

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 16:17:35 ; Search time 71.77 Seconds
(Without alignments)
118.689 Million cell updates/sec

Title: US-09-052-089a-4
Perfect score: 1075
Sequence: 1 KTIINKLPFDLAQEEENVLD.....DLQADQETSLRKSDPP 220

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	170	15.8	3210	1	CENP_HUMAN
2	169.5	15.8	1957	1	YD86_SCHPO
3	169	15.7	1976	1	MYHA_BOVIN
4	168.5	15.7	1325	1	G160_MOUSE
5	167.5	15.6	1203	1	XCPC_XENLA
6	166	15.4	2017	1	MTSN_DROME
7	163	15.2	1790	1	USO1_YEAST
8	162.5	15.1	2116	1	MYSD_DICDI
9	162	15.1	895	1	RA50_THEVO
10	162	15.1	1938	1	MYA_AEOIR
11	160.5	14.9	1744	1	TANA_XENLA
12	160.5	14.9	1966	1	MTSB_CABEL
13	160	14.9	1972	1	MYHB_MOUSE
14	159	14.8	963	1	KINH_HUMAN
15	159	14.8	1972	1	MYHB_HUMAN
16	159	14.8	1972	1	MYHB_RABIT
17	157.5	14.7	886	1	RA50_ARCFU
18	156.5	14.6	1509	1	MTSN_ACACA
19	155	14.4	962	1	VDP_HUMAN
20	154.5	14.4	2411	1	MYSA_DROME
21	154	14.3	896	1	EP15_HUMAN
22	154	14.3	1976	1	MYHA_HUMAN
23	153.5	14.3	1938	1	MYSD_CABEL
24	153	14.2	1093	1	TMFL_HUMAN
25	153	14.2	1938	1	MYHD_HUMAN
26	152	14.1	1969	1	MYSA_CABEL
27	151.5	14.1	995	1	HIP1_HUMAN
28	151	14.0	880	1	RA50_PYPAB
29	150.5	14.0	882	1	MYSP_CABEL
30	150	14.0	976	1	SCPL_HUMAN
31	150	14.0	1937	1	MYHB_HUMAN
32	150	14.0	1961	1	MYH9_RAT
33	149	13.9	1935	1	MYSS_CYPCA

34	149	13.9	1960	1	MYH9_HUMAN	P35579 homo sapien
35	149	13.9	1978	1	MYHB_CHICK	P10587 gallus gall
36	148.5	13.8	501	1	MYSD_RABIT	O99105 cryptotagus
37	148.5	13.8	727	1	MFPL_ARATH	O91485 arabidopsis
38	148.5	13.8	1290	1	XCPC_XENLA	P50532 xenopus lae
39	148.5	13.8	1427	1	REST_HUMAN	P30622 homo sapien
40	148.5	13.8	1934	1	MYH7_MESAU	P13540 mesocricetu
41	148.5	13.8	1935	1	MYH7_HUMAN	P12883 homo sapien
42	148.5	13.8	1935	1	MYH7_PIG	P79293 sus scrofa
43	148	13.8	1102	1	MYSC_CHICK	P29616 gallus gall
44	147.5	13.7	579	1	G160_HUMAN	O08378 homo sapien
45	147.5	13.7	880	1	MYSP_BRUMA	O01202 brugia mala

ALIGNMENTS

RESULT 1
ID CENP_HUMAN STANDARD: PRT: 3210 AA.
AC P49454: Q13246: Q13171:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CENP-F kinetochore protein (centromere protein F) (Mitosis) (AH antigen).
GN CENP-F.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast carcinoma;
RX MEDLINE=95348175; PubMed=7542657;
RA Liao H., Winkfein R.J., Mack G., Ratner J.B., Yen T.J.;
RT "CENP-F is a protein of the nuclear matrix that assembles onto kinetochores at late G2 and is rapidly degraded after mitosis.";
RL J. Cell Biol. 130:507-518(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95379848; PubMed=7651420;
RA Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B., Jones D., Yang-Feng T.L., Lee W.-H.;
RT "Characterization of a novel 350-kilodalton nuclear phosphoprotein that is specifically involved in mitotic-phase progression.";
RL Mol. Cell. Biol. 15:5017-5029(1995).
RN [3]
RP SEQUENCE OF 2194-3210 FROM N.A.
RX MEDLINE=95336446; PubMed=7612011;
RA Li Q., Ke Y., Kapp J.A., Fertig N., Medsger T.A. Jr., Joshi H.C.;
RT "A novel cell-cycle-dependent 350-kDa nuclear protein: C-terminal domain sufficient for nuclear localization.";
RL Biochem. Biophys. Res. Commun. 212:220-228(1995).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=95370296; PubMed=7642639;
RA Zhu X., Chang K.-H., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;
RT "The C terminus of mitotin is essential for its nuclear localization, centromere/kinetochore targeting, and dimerization.";
RL J. Biol. Chem. 270:19545-19550(1995).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=96437347; PubMed=9763420;
RA Chan G.K.T., Schaar B.T., Yen T.J.;
RT "Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hBUB1.";
RL J. Cell Biol. 143:49-63(1998).
CC -1- FUNCTION: PROBABLY REQUIRED FOR KINETOCORE FUNCTION, INVOLVED IN CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH RTINOBLASTOMA PROTEIN (RB), CENP-E AND BUB1.
CC -1- SUBUNIT: HOMO- OR HETERODIMER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS),

```

CC REORGANIZATION TO THE KINETOCORE/CENTROMERE (CORONAL SURFACE OF
CC THE OUTER PLATE) AND THE SPINDLE DURING MITOSIS.
CC -1- DEVELOPMENTAL STAGE: GRADUALLY ACCUMULATES DURING THE CELL CYCLE.
CC -1- PTM: HYPERPHOSPHORYLATED DURING MITOSIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U19769; AAA82889.1; -
DR EMBL; U30872; AAA82935.1; -
DR EMBL; U25725; AAA86889.1; -
DR HSP; P02649; ILE4.
DR MIM; 600236; -.
KW Chromosomal protein; Nuclear protein; Centromere; Coiled coil;
KW Mitosis; Phosphorylation; Antigen; Cell cycle; Repeat.
FT DOMAIN 14 197
FT DOMAIN 273 769
FT DOMAIN 823 1328
FT DOMAIN 1642 1746
FT DOMAIN 1862 2987
FT DOMAIN 2207 2568
FT REPEAT 2207 2386
FT REPEAT 2389 2568
FT REPEAT 3015 3032
FT CONFLICT 16 16
FT CONFLICT 250 250
FT CONFLICT 272 272
FT CONFLICT 611 611
FT CONFLICT 1494 1589
FT CONFLICT 1611 1611
FT CONFLICT 1811 1811
FT CONFLICT 2242 2243
FT CONFLICT 2335 2335
FT CONFLICT 2492 2492
FT CONFLICT 2545 2561
SQ SEQUENCE 3210 AA; 367589 MW; 11D83324960E4334 CRC64;

Query Match 15.8%; Score 170; DB 1; Length 3210;
Best Local Similarity 24.3%; Pred. No. 0.039;
Matches 69; Conservative 51; Mismatches 96; Indels 68; Gaps 9;

OY 1 KTINKLFEDLAQEEEN---VLDAEFLKNEIDSVKAO-----SOK 38
DB 2249 KDKVENLEERELQMESEENQELVILDAENSKAEVETLKTQIEEMARSLKVFELDLVTLRSEK 2308
OY 39 D-----REKDSQAIIIDTLRDTL-----ERRNATVESLQNALNKAEMLC 77
DB 2309 ENLTQKQIOEKQGLSELDKLLSKSLLEKEQAEIQIKESKTAVENTLQNLKELNEAV 2368
OY 78 STL---KKOMKFLBOROD---FTKQAREEFAHRLCKCKMTQETLELLOSORSE----- 124
DB 2369 AALCGDQEIEMKATQSDPREEHQRLNSTEKLRLARLEADEKQOLCVLQQLKSEHHAD 2428
OY 125 -----VEEMIRDMGVGSAVQOLAVCVSLKKEYENLKERRATGELADRLKKDLVSSRS 179
DB 2429 LTKRVEVLELERELIARINQCHALAEVNSKGEVETLKAKLEGMTQSLRGELDVVTIRS 2488
OY 180 KLTKLNTLELDQ-----AKLEL--RSAOKDLOSADQETLSLRKKS 216
DB 2489 EKEDLTNELQKEQERISELETLINSSFEVTLQEKBOEKVMKEKS 2532

RESULT 2
YD86_SCHPO STANDARD; PRT; 1957 AA.
AC Q10411;

```

```

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 222.8 kDa protein Clf3.06C in chromosome 1.
GN SPAC1f3.06C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z70690; CA94624.1; -
DR Hypothetical protein.
SQ SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

```

```

Query Match 15.8%; Score 169.5; DB 1; Length 1957;
Best Local Similarity 22.4%; Pred. No. 0.025;
Matches 54; Conservative 54; Mismatches 96; Indels 37; Gaps 4;

OY 15 EENVLDAEFLKNEIDSVKAOQLSQDKREKDSQAIIIDTLRDTLEERNATVESLQNALNKAE 74
DB 1386 EDNOLATNKKLKNQDLHNLQETRLKEDVLEKESLISLESLSQROKESLSLDAKKELE 1445
OY 75 -MLSTLKQKPFLEORO-----DETQAREEFAHRLCKCKMTQETLELLOSORSEVE 126
DB 1446 HMLDDTSKNSLSMEKIESINSSLDKSFELASAVEKIGALQKLSHESLSMEIKSOLQ 1505
OY 127 EMIRDMGVGSAVQOLAVCVSLKKEYEN-----LKEARRATG 164
DB 1506 EAKKEIGVDESTIOELDHEITASKNNEGKLNKDSIIRDSENTLEQNLNLAEKSAVK 1565
OY 165 ELARLKKDLVSSRSKLT-----NTELDQAKLELRSAQKQDQSDQETLSLRKKS 217
DB 1566 RLSTERSESLIQNSRLADLEYHKQSVESSEIGRSKLASTTEBLQLAENRSLSTTRML 1625
OY 218 D 218
DB 1626 D 1626

RESULT 3
MYHA_BOVIN STANDARD; PRT; 1976 AA.
AC Q27991;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, nonmuscle type B (cellular myosin heavy chain,
DE type B) (Nonmuscle myosin heavy chain-B) (NMHC-B).
GN MYH10.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Ohara M., Ishiguro N., Shinaawa M.;
RT "Bos taurus nonmuscle myosin heavy chain B mRNA, complete cds."

```

Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

123 SEVENIRDMGVQSAVEQLAVYCVSLKREYEN-----LKEARKATG----- 164

1258 DAQVQELHAKVSGDRULVLEAKRANKQNELNVSTLLEARKKKGIKFAKDAAGLESL 1317

165 -----ELADLKKDDVSSRSKLTNTLTDQALFLRSADKQDSADQEI 209

1318 QDTQELLQEEFTROKLNLSIRIQ-LEERSSIQEQOEEEBEAR--RSLKQLQALQAL 1373

210 TSLRKSD 218

1374 TDTKKKYDD 1382

RESULT 4

G160 MOUSE STANDARD; PRT; 1325 AA.

AC P55937;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Golgin-160 (Male-enhanced antigen-2) (MEA-2).

GN GOLG3 OR MEA2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CD-1; TISSUE=Testis;

RX MEDLINE=97217683; PubMed=9063644;

RA Kondo M., Sutou S.;

RT Cloning and molecular characterization of cDNA encoding a mouse

RT male-enhanced antigen-2 (Mea-2): a putative family of the Golgi

RT autoantigen.";

RL DNA Seq. 7:71-82(1997).

CC - FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR

CC TESTIS DEVELOPMENT. PROBABLY IDENTICAL WITH THE SEROLOGICALLY

CC DETECTABLE MALE ANTIGEN (SDM).

CC - TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE

CC FOUND IN SPERMATIDS DURING SPERMATOGENESIS. NO EXPRESSION IN

CC LEYDIG CELLS, SPERMATOGENIA, OR SPERMATOCYTES.

CC - SIMILARITY: HIGH, TO HUMAN GOLGIN-160.

CC - CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-19 OR MET-30 IS THE

CC INITIATOR.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AB022023; BAA36494.1; -

DR EMBL: 015716; AAA87715.1; -

DR HSSP: P08799; ILVK.

DR InterPro: IPR000048; IQ.

DR InterPro: IPR004009; Myosin_N.

DR InterPro: IPR002928; Myosin_tail.

DR InterPro: IPR002017; Spectrin.

DR InterPro: IPR001609; myosin_head.

DR Pfam: PF00612; IQ; 1.

DR Pfam: PF00063; myosin_head; 1.

DR Pfam: PF02736; Myosin_N; 1.

DR Pfam: PF01576; Myosin_tail; 1.

DR PRINTS: PR00193; MYOSINHEAVY.

DR ProDom: PD000355; myosin_head; 1.

DR SMART: SM00015; IQ; 1.

DR SMART: SM00242; MYSC; 1.

DR PROSITE: PS50096; IQ; 1.

DR Myosin; ATP-binding; Calmodulin-binding; Actin-binding;

DR Coiled coil; Alkylation; Multigene family.

KW MYOSIN HEAD-LIKE.

FT DOMAIN 1 785

FT DOMAIN 786 815

FT DOMAIN 845 1976

FT NP_BIND 178 185

FT MOD_RES 701 701

FT MOD_RES 711 711

SEQUENCE 1976 AA; 229097 MW; 61435451C0F790 CRC64;

Query Match 15.7%; Score 169; DB 1; Length 1976;

Best Local Similarity 21.4%; Pred. No. 0.027;

Matches 66; Conservative 58; Mismatches 87; Indels 98; Gaps 10;

4 INKLFDLAGEENV-----LDAEF--LNKELDSVKAQLSOK 38

1078 IDELKIYVAKKEELQALRGDEFTLHKNNALKVYRELQALQELQDESESEKASNNKA 1137

39 DREKRSQAIIIDTLRDTLEERNATVESLQNALNKAEMLCSTLKK-----QMKFL 87

1138 EKQKRDLSEELKALKTLELTPTAAQCELRTRKEQEVLELKALBEETKSHQAQIDQM 1197

88 EQHQ-----DETKQAREFAHLK-----CKMKTMOIELLLQSOR--- 122

1198 RQKHATLLELSEQLQAKRFRKANLEKNKQGLTQNKELACEVAVLQOVKAESSEHKRKL 1257

123 SEVENIRDMGVQSAVEQLAVYCVSLKREYEN-----LKEARKATG----- 164

1258 DAQVQELHAKVSGDRULVLEAKRANKQNELNVSTLLEARKKKGIKFAKDAAGLESL 1317

165 -----ELADLKKDDVSSRSKLTNTLTDQALFLRSADKQDSADQEI 209

1318 QDTQELLQEEFTROKLNLSIRIQ-LEERSSIQEQOEEEBEAR--RSLKQLQALQAL 1373

210 TSLRKSD 218

1374 TDTKKKYDD 1382

RESULT 4

G160 MOUSE STANDARD; PRT; 1325 AA.

AC P55937;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Golgin-160 (Male-enhanced antigen-2) (MEA-2).

GN GOLG3 OR MEA2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CD-1; TISSUE=Testis;

RX MEDLINE=97217683; PubMed=9063644;

RA Kondo M., Sutou S.;

RT Cloning and molecular characterization of cDNA encoding a mouse

RT male-enhanced antigen-2 (Mea-2): a putative family of the Golgi

RT autoantigen.";

RL DNA Seq. 7:71-82(1997).

CC - FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR

CC TESTIS DEVELOPMENT. PROBABLY IDENTICAL WITH THE SEROLOGICALLY

CC DETECTABLE MALE ANTIGEN (SDM).

CC - TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE

CC FOUND IN SPERMATIDS DURING SPERMATOGENESIS. NO EXPRESSION IN

CC LEYDIG CELLS, SPERMATOGENIA, OR SPERMATOCYTES.

CC - SIMILARITY: HIGH, TO HUMAN GOLGIN-160.

CC - CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-19 OR MET-30 IS THE

CC INITIATOR.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: D78270; BAA19612.1; -

DR HSSP: P18852; ISCG.

DR MGD; MGI:96958; GOL9a3.

KW Spermatogenesis; Developmental protein.

FT DOMAIN 201 204

SEQUENCE 1325 AA; 149880 MW; 3230656962C687B0 CRC64;

Query Match 15.7%; Score 168.5; DB 1; Length 1325;

Best Local Similarity 23.8%; Pred. No. 0.019;

Matches 69; Conservative 50; Mismatches 84; Indels 87; Gaps 12;

10 DLAGEENVLDA-EFLNE-----LDSVKAQLSOKDREKR-----DSQAIIIDTLRD 54

590 ELQREADSRDAIHFILQNEKIVLEVALQSAKSEEDRGARRLEEDTETSGLLEQLRQ 649

55 TLBERNATVESLQNALNKAEMLCSTLKKOM-----KFLQD-----RODET--- 96

650 DLAVKSNQVEHLOE-----TATLRKQKQKVEQFQOVKVVAVYRQATSKDQLINE 702

```

0Y 97 AREENHRLKCKMKTMEQLELLLOSORSVE-----EMIRDMGVGSAVEQULAVCVSLKKE 1522
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 703 LKATKRRLDSEMKELROELIKLOEKKTVESHSLRDKDSVLVHQVAAELEGHLOSVOKE 7623
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
0Y 153 YEN-----LKEA-----RKATGSLARLTKDLVSSRSKLTINTLELDOA 191
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 763 RDMEIHLLOSLKPEKQWIALTEANETLKKOIELLOEARKKATIEOKKMKKRLGSLDTLSA 8222
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
0Y 192 KLELSAQKDLQSA-----DQEIISLRKK-----SDPP 219
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 823 QKEKMTTKKAVENAVSILSRRLDALASKATYDDELNOLRQSTGGSSDP 872
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

```

RESULT 5
XCPE_XENLA STANDARD: PRT: 1203 AA.
AC P50533.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Chromosome assembly protein XCAP-E.
GN XCAP-E.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95042742; PubMed=7954811;
RA Hirano T., Mitchison T.J.;
RT "A heterodimeric coiled-coil protein required for mitotic chromosome
  condensation in vitro.";
RL Cell 79:449-458(1994).
CC -1- FUNCTION: REQUIRED FOR BOTH ASSEMBLY AND STRUCTURAL MAINTENANCE OF
  MITOTIC CHROMOSOMES.
CC -1- SUBUNIT: ASSOCIATES WITH XCAP-C PROBABLY AS HETERODIMER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. DURING CHROMOSOME ASSEMBLY IN
  MITOTIC EXTRACTS, XCAP-C/E WAS RECRUITED TO THE CHROMATIN AND
  FORMED A DISCRETE INTERNAL STRUCTURE WITHIN ASSEMBLED CHROMOSOMES.
CC -1- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS
  FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U13674; AAA64680.1; -.
DR HSSP: P07751; IAN3.
DR InterPro: IPR003405; SMC_C.
DR InterPro: IPR003395; SMC_N.
DR Pfam: PF02483; SMC_N; 1.
DR Pfam: PF02463; SMC_N; 1.
KW Mitosis; ATP-binding; Coiled coil; Nuclear protein.
FT NP_BIND 32 39 ATP (POTENTIAL).
FT DOMAIN 172 356 COILED COIL (POTENTIAL).
FT DOMAIN 397 513 COILED COIL (POTENTIAL).
FT DOMAIN 670 1032 COILED COIL (POTENTIAL).
FT DOMAIN 1095 1121 ALA/ASP-RICH (DA-BOX).
SO SEQUENCE 1203 AA; 136341 MW; 04323DD0027DF309 CRC64;

Query Match 15.6%; Score 167.5; DB 1; Length 1203;
Best Local Similarity 24.6%; Pred. No. 0.02;
Matches 51; Conservative 45; Mismatches 80; Indels 31; Gaps 5;

```

```

Db 721 EKMSEBLLLOTKLOQSSYHHKQOE--LDLSLKQIIESEBFLKMTKVOYKAEKEFVLE 778
Qy 82 KQMEFL-ORODETRQANE-----EAHRLCKMKMTQIEBLLLOSORSEVEEMIRD 131
Db 779 HKMNAEAEERERBELKEAQQLDPAKKKADASNNKMKKEQOVDALVL-----ELEELKRE 833
Qy 132 MGVGOSAEVOLAVYCVSLKKEYENLKEKRRKTTGELADRLKDLVSSSKLTKLTLELDOA 191
Db 834 Q-----TTYKQOIEVDEEMKAYQOADSMSAEVSKNKEAVKKAODELAKO 879
Qy 192 KLELSAQKDLQSDAQDEITSRRKSSD 218
Db 880 KEIITGHDKETIKTSSEAGKLRNNND 906

```

RESULT	6			
MYSN_DROME		STANDARD:	PRF:	2017 AA.
ID	MYSN_DROME			
AC	Q99323:			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Myosin heavy chain, non-muscle (zipper protein) (Myosin I1).			
GN	zip.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Embryo;			
RX	MEDLINE=90349606; PubMed=2117279;			
RA	Ketchum A.S., Stewart C.T., Stewart M., Kiehart D.P.;			
RT	"Complete sequence of the Drosophila nonmuscle myosin heavy-chain			
RT	transcript: conserved sequences in the myosin tail and differential			
RT	splicing in the 5' untranslated sequence."			
RL	Proc. Natl. Acad. Sci. U.S.A. 87:6316-6320(1990).			
CC	-1- FUNCTION: NONMUSCLE MYOSIN APPEARS TO BE RESPONSIBLE FOR			
CC	CYTOPLASMIC CONTRACTILITY. REQUIRED FOR MORPHOGENESIS AND CYTOKINESIS.			
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A			
CC	SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 1 IQ DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/or-send-an-email-to-licence@isb-sib.ch).			
CC	-----			
CC	EMBL; M35012; AAA28713.1; -.			
DR	PIR; A36014; A36014.			
DR	PIR; B36014; B36014.			
DR	HSSP; P08799; 1MMD.			
DR	FLYBase; FBgn0005634; zip.			
DR	InterPro; IPR0000048; IQ.			
DR	InterPro; IPR004009; Myosin_N.			
DR	InterPro; IPR002928; Myosin_tail.			
DR	InterPro; IPR001609; myosin_head.			
DR	Pfam; PF00612; IQ: 1.			
DR	Pfam; PF00063; myosin_head; 1.			
DR	Pfam; PF02736; Myosin_N; 1.			
DR	Pfam; PF01576; Myosin_tail; 1.			
DR	PRINTS; PR00193; MYOSINHEAVY.			
DR	ProDom; PD000355; myosin_head; 1.			
DR	SMART; SM00015; IQ: 1.			
DR	SMART; SM00242; MYSC; 1.			
DR	PROSITE; PS50096; IQ: 1.			
DR	Myosin; Alternative splicing; Coiled coil; Actin-binding;			
DR	ATP-binding; Calmodulin-binding.			

```
FT DOMAIN 1 829 MYOSIN HEAD-LIKE.
FT DOMAIN 830 859 IQ.
FT DOMAIN 866 2017 COILED COIL (POTENTIAL).
FT NP_BIND 225 232 ATP.
FT DOMAIN 250 260 25 KDA/50 KDA JUNCTION.
FT DOMAIN 682 694 50 KDA/20 KDA JUNCTION.
FT DOMAIN 705 727 ACTIN-BINDING.
FT DOMAIN 742 758 REACTIVE SULFHYDRYL/ACTIN-BINDING.
FT DOMAIN 1303 2017 LIGHT MEROMYOSIN (LMM).
FT DOMAIN 1303 1970 ALPHA-HELICAL TAILPIECE (LMM).
FT DOMAIN 1971 2017 GLOBULAR TAILPIECE.
FT VARSPLIC 1 45 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 2017 AA; 232016 MW; 73E3CB02BA8F2528 CRC64;

Query Match 15.4%; Score 166; DB 1; Length 2017;
Best Local Similarity 25.2%; Pred. No. 0.04; Mismatches 96; Indels 38; Gaps 8;
Matches 61; Conservative 47;

OY 1 KTIINKLFEDLAOEENVIDAE--FLKNEDSVKAOISQKREKRDQSO-----AI 48
DB 1263 KYLKEK-----AKGTLEAENADLTELRSVNSSROENDRKKAQESQIAELQYKLA 1314
OY 49 IDTLRDTLEER---NATVESIQNALNKAEMLCSTLKKQKLFQRODETQKAREAHN- 103
DB 1315 IERARSELQEKCTKLOQEAENITNOLEAEIKASAAVKSASNNESQLEAQLLEETRQ 1374
OY 104 ---LKKMKMTMEQIELLOSQSEVEEMTRMGVGSAAVQALVYCVSLKKEYNLKEAR 160
DB 1375 KIGLSKRLROIESEKENLOQLDEDEAKRNY--EKLEAVTTONQEIKKKAEDADLA 1431
OY 161 KATGELADRLKDLVSSRSKLTU--NPELDAQLELRSQKD---LQSAQDEITSRL 213
DB 1432 KELEBEKKRLNKQIEALERVKELIQAQNDRLDKSKKQISQELDEDTLELAQRTKYLE 1491
OY 214 KK 215
DB 1492 KK 1493

RESULT 7
USOL_YEAST STANDARD: PRT; 1790 AA.
AC P25386;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN Intracellular protein transport protein USOL.
OS USOL OR INT1 OR YDL058M.
OC Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetiales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RX MEDLINE=91185402; PubMed=2010462;
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA Yamasaki M.;
RT "A cytoskeleton-related gene, usol, is required for intracellular
RT protein transport in Saccharomyces cerevisiae.";
RL J. Cell Biol. 113:245-260(1991).
RN [12]
RP SEQUENCE OF 782-1790 FROM N.A.
RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA Kendrick K.E.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [13]
RP SEQUENCE OF 1-8 FROM N.A.
RA Bai Y., Symington L.S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC COMPLEX.
```

```
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC ER AND THE GOLGI COMPLEX.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
CC -----
DR EMBL: X54378; CA38253.1; -
DR EMBL: L03188; AAB00143.1; -
DR EMBL: U53668; AAB6659.1; -
DR PIR: A38455; A38455.
DR HSSP: P80220; 1DIP.
DR SCD: S00002216; USOL.
DR InterPro: IPR002017; Spectrin.
KW Transport; Protein transport; Golgi stack; Cytoskeleton; coiled coil.
FT DOMAIN 1 724 GLOBULAR HEAD.
FT DOMAIN 725 1790 COILED COIL (POTENTIAL).
FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).
FT DOMAIN 921 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.
FT DOMAIN 1172 1786 ASF/GLU-RICH (ACIDIC).
FT CONFLICT 847 847 G->E (IN REF. 2).
FT CONFLICT 924 924 E->K (IN REF. 2).
FT CONFLICT 1253 1253 V->I (IN REF. 2).
FT CONFLICT 1319 1319 I->V (IN REF. 2).
FT CONFLICT 1461 1461 N->S (IN REF. 2).
FT CONFLICT 1581 1581 G->S (IN REF. 2).
FT CONFLICT 1600 1600 I->V (IN REF. 2).
FT CONFLICT 1661 1661 R->S (IN REF. 2).
FT CONFLICT 1772 1772 D->DEEDDEE (IN REF. 2).
SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9F04818 CRC64;

Query Match 15.2%; Score 163; DB 1; Length 1790;
Best Local Similarity 23.7%; Pred. No. 0.052; Mismatches 77; Indels 30; Gaps 7;
Matches 51; Conservative 57;

OY 10 DLAOEENVIDAEFLKNEELDSVKAQISQKREKRDQAIITDLTLEERNATVESIQNA 69
DB 1424 ELLEEKONTIKS--LODEI-----LSYKDKITRNDEKLTLSIERDNKRD---LESLEQ 1471
OY 70 LNKAEMLCSTLKKQKMFLEQRODETQKAREAHNLCMKMTMEQIELLOSQSEVEEMI 129
DB 1472 LRAQESKAVVERGLKLEBESSKEKAELEKSKEM-----MKKLESTISNETELKSSM 1525
OY 130 RDMGVGSAAVEOLAIVYCVSLKKEYNLKEARKATGELADRL---KDLVSSRSKLT--T 183
DB 1526 ETIRKSDEKLEQ-----SKKSAEEDIKNQLQHEKSDLSIRNESEKQIEELKSLRIEAK 1579
OY 184 LNTLEDAQLELNSAQKLOLSAQDOEITSLRKSSD 218
DB 1580 SGSELETVKQELNNQAQEKIRINAEENTVLSKSLIED 1614

RESULT 8
MYS2_DICDI STANDARD: PRT; 2116 AA.
AC P08799;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin II heavy chain, non muscle.
GN MHCA.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
```

ON NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87092266; Pubmed=3540939;
 RA Marlick H.M., de Lozanne A., Jeanvond L.A., Spudich J.A.;
 RT "Conserved protein domains in a myosin heavy chain gene from
 R7 Dictyostelium discoideum.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
 RN [2]
 RP PHOSPHORYLATION SITES, AND MUTAGENESIS.
 RC STRAIN-AX2;
 RX MEDLINE=90353583; Pubmed=2387408;
 RA Lueck-Vielmeier D., Schlachter M., Grabatin B., Wippler J.,
 RA Gertsch G.;
 RT "Replacement of threonine residues by serine and alanine in a
 R7 phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
 RL FEBS Lett. 269:239-243(1990).
 RN [3]
 RP PHOSPHORYLATION SITES.
 RX MEDLINE=88112226; Pubmed=2828113;
 RA Wagle G., Noegel A., Scheel J., Gertsch G.;
 RT "Phosphorylation of threonine residues on cloned fragments of the
 R7 Dictyostelium myosin heavy chain.";
 RL FEBS Lett. 227:71-75(1988).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
 RX MEDLINE=95345066; Pubmed=7619795;
 RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
 RA Rayment I.;
 RT "X-ray structures of the myosin motor domain of Dictyostelium
 R7 discoidium complexed with MgADP. Bex and MgADP.ALf4-.";
 RL Biochemistry 34:8960-8972(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
 RX MEDLINE=95345067; Pubmed=7619796;
 RA Smith C.A., Rayment I.;
 RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
 R7 truncated head of Dictyostelium discoidium myosin to 2.7-A
 resolution.";
 RL Biochemistry 34:8973-8981(1995).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RX MEDLINE=96206189; Pubmed=8611530;
 RA Smith C.A., Rayment I.;
 RT "X-ray structure of the magnesium(II)-ADP.vanadate complex of the
 R7 Dictyostelium discoidium myosin motor domain to 1.9-A resolution.";
 RL Biochemistry 35:5404-5417(1996).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
 RX MEDLINE=97452580; Pubmed=9305951;
 RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
 RT "X-ray structures of the MgADP, MgATPgmasS, and MgAMPNP complexes
 R7 of the Dictyostelium discoidium myosin motor domain.";
 RL Biochemistry 36:11619-11628(1997).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RX MEDLINE=98070605; Pubmed=9405148;
 RA Bauer C.B., Kuhlman P.A., Baasch C.R., Rayment I.;
 RT "X-ray crystal structure and solution fluorescence characterization
 R7 of Mg₂(3')-O-(N-methylanthraniloyl) nucleotides bound to the
 Dictyostelium discoidium myosin motor domain.";
 RL J. Mol. Biol. 274:394-407(1997).
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
 CC ACTIVITY THAT IS ACTIVATED BY ACTIN.
 CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
 CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
 CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
 CC (MLC-2).
 CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
 CC CORTEX.
 CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMH) AND 1 HEAVY MEROMYOSIN (HMH). IT CAN BE FURTHER
 CC SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED

```

CC SUBFRAGMENT (S2).
CC -I- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -I- PPM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.
CC -I- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE
CC ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
CC POSITION (688).
CC -I- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL, M14628; AAA33227.1; -.
CC DR PIR: A26655; A26655.
CC DR PIR: S00250; S00250.
CC DR PDB: 1MMA: 03-DEC-97.
CC DR PDB: 1MMD: 17-AUG-96.
CC DR PDB: 1MMG: 03-DEC-97.
CC DR PDB: 1MMN: 03-DEC-97.
CC DR PDB: 1MND: 17-AUG-96.
CC DR PDB: 1MNE: 17-AUG-96.
CC DR PDB: 1VOM: 23-DEC-96.
CC DR PDB: 1VVK: 28-JAN-98.
CC DR DictyDB: DD01008; mhca.
CC DR InterPro: IPR000048; IQ.
CC DR InterPro: IPR004009; Myosin_N.
CC DR InterPro: IPR001609; myosin_head.
CC DR Pfam: PF00612; IQ_2.
CC DR Pfam: PF00063; myosin_head; 1.
CC DR Pfam: PF02736; Myosin_N; 1.
CC DR PRINTS: PR00193; MYOSINHEAVY.
CC DR ProDom: PD000355; myosin_head; 1.
CC DR SMART: SM00015; IQ_1.
CC DR SMART: SM00242; MYSC; 1.
CC DR PROSITE: PS00096; IQ_1.
CC KW Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;
CC Calmodulin-binding; Methylation; Alkylation; Phosphorylation.
CC FT DOMAIN 1 761
CC FT DOMAIN 762 791
CC FT NP_BIND 817 2116 COILED COIL (POTENTIAL).
CC FT NP_BIND 179 186 ATP.
CC FT DOMAIN 638 660 ACTIN-BINDING.
CC FT DOMAIN 738 752 ACTIN-BINDING.
CC FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).
CC FT MOD_RES 678 678 METHYLATION (SH-1).
CC FT MOD_RES 1823 1823 PHOSPHORYLATION (BY MHCK).
CC FT MOD_RES 1833 1833 PHOSPHORYLATION (BY MHCK).
CC FT MOD_RES 2029 2029 PHOSPHORYLATION (BY MHCK).
CC SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770B1BE56A1 CRC64;

Query Match 15.1%; Score 162.5; DB 1; Length 2116;
Best Local Similarity 25.0%; Pred. No. 0.066;
Matches 67; Conservative 52; Mismatches 86; Indels 63; Gaps 11.

5 NKLFPDLAEGEENVDAE-FLKNEIDSVKQSLQSKD---REKRDQALIDTLRDPLEERN 60
:||||| : ||||| : || :||| : ||| :||| :||| :||| :||| :|||
Db 852 DLEKSLKQDTEBNVLDLQRLQKLAKEETLKAMTDSKDALEAQNRELEIRYEDNESELDKK 911
:||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Oy 61 ATVESLON-----ALKKAKEMLCSTLKKOMKFLFRODQETKQAREEAHRLKC 106
:||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db 912 LALNELQNGKRSVEEKVNRLEELQDEQKLRNTLEKIKKKYEELEELDKMKRVADG----- 965
:||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Oy 107 KMKTKQELILLQSORSEVEEMI-----RDMGV-----GQSAVDELAVYCVS---- 148
:||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||

```

```

Db 966 QSDTISRLKIKDELOVEELTESFSESKDKGLEKTRVRLQSELDITVRLDSETKD 1025
QY 149 -----LKEEYENLKAEKRA-TGELADRLKDLVSRSLKLTNLELDO-----AKL 193
Db 1026 KSELLRKKKLEELKQVDELAETRAKLAOE--AAKKKLQGEYTELNEKFNSEVTARS 1083
QY 194 ELRSQKDLQS---ADQETISLRKSD 217
Db 1084 NVEKSKKLESQLVAVNNELDEKKKND 1111

RESULT 9
RA50_THEVO
ID RA50_THEVO STANDARD; PRT; 895 AA.
AC P58302;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE DNA double-strand break repair rad50 ATPase.
GN RAD50 OR TVG0235331.
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaaceae;
OC Thermoplasma.
OX NCBI_TaxID=50339;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRAIN-GSSI / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA Ninochiba T., Yamamoto Y., Araiaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of Thermoplasma volcanium.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mre11 complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (By
CC similarity).
CC -1- SUBUNIT: Forms a complex with mre11 (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUPERFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AP000991; BAB59370.1; -
DR InterPro: IPR003439; ABC_Transport.
DR InterPro: IPR000879; Cecropin.
DR InterPro: IPR001238; RecF.
DR InterPro: IPR003395; SMC_N.
DR Pfam: PF00470; RecF. 1.
DR Pfam: PF02463; SMC_N. 1.
KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
FT NP_BIND 30
FT DOMAIN 171 737 COILED COIL (POTENTIAL).
SQ SEQUENCE 895 AA; 103223 MW; 4A47DA9287E82D3A CRC64;

Query Match 15.1%; Score 162; DB 1; Length 895;
Best Local Similarity 24.3%; Pred. No. 0.029;
Matches 52; Conservative 52; Mismatches 80; Indels 30; Gaps 6;

```

```

Db 227 INIKENKRD--HLNEELHRLNMQLEFIKKYKMELEASQSRKASIEMEYVKLSIEBELK 283
QY 141 QLAIVCYSLTK-----EVENLKEARKATGELADRLK-----KLVSRSLK-- 161
Db 284 RLNNAAVVRNKEIIEINIKKDLQSLEIEIGKLSKYDEAHRKLEDLQSFSEFLE 343
QY 182 -KTLNLELDAQLELRSAQKDLQSADEITSLRK 214
Db 344 KKKRKEDLDKRLSKLEDEDNQSAVNIENIKK 377

RESULT 10
MYS_AEOIR
ID MYS_AEOIR STANDARD; PRT; 1938 AA.
AC P24733;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, striated muscle.
DE Myosin heavy chain, striated muscle.
OS Aequipecten irradians (Bay scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea;
OC Pectinoidea; Pectinidae; Argopecten.
OX NCBI_TaxID=31199;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=92011595; PubMed=1917970;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Complete primary structure of a scallop striated muscle myosin heavy
RT chain. Sequence comparison with other heavy chains reveals regions
RT that might be critical for regulation.";
RL J. Biol. Chem. 266:18469-18476(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=91088319; PubMed=2263488;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Nucleotide sequence of full length cDNA for a scallop striated
RT muscle myosin heavy chain.";
RL Nucleic Acids Res. 18:7158-7158(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
RX MEDLINE=94173332; PubMed=8127365;
RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,
RA Szent-Gyorgyi A.G., Cohen C.;
RT "Structure of the regulatory domain of scallop myosin at 2.8-A
RT resolution.";
RL Nature 368:306-312(1994).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.
RX MEDLINE=96419133; PubMed=8805510;
RA Houdusse A., Cohen C.;
RT "Structure of the regulatory domain of scallop myosin at 2-A
RT resolution: implications for regulation.";
RL Structure 4:21-32(1996).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```



```
CC EMBL: X55714; CAA39247.1; -
DR PIR: S13557; S13557.
DR PIR: A40997; A40997.
DR PDB: 1SCM; 30-APR-94.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KM Myosin: Muscle protein: Coiled coil: Thick filament: Actin-binding;
KW ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.
FT DOMAIN 1 777
FT DOMAIN 778 805
FT DOMAIN 836 1938
FT DOMAIN 836 1938
FT NP_BIND 176 183
FT MOD_RES 703 703
FT MOD_RES 778 821
FT HELIX 822 823
FT TURN 825 833
FT TURN 834 835
SQ SEQUENCE 1938 AA; 222821 MW; A5CCEA127D1A4896 CRC64;
```

```
Query Match 15.1%; Score 162; DB 1; Length 1938;
Best Local Similarity 24.5%; Pred. No. 0.064;
Matches 73; Conservative 50; Mismatches 89; Indels 86; Gaps 12;
```

```
OY 1 KTIINKLFFD-----LAQEEENVLDAEFLKNELDVSKAQL-----SOK 38
DB 904 KLIMQKADFESQIKLEBERLDEDAADLEGIKKMEADANAKKIDGLDLENTLQKAEQ 963
OY 39 DREKRDQAIIDTL-----RDILEERN-ATVESIQ-----NALKKA 74
DB 964 DKAKKDNO--ISTIQGELISQODEHIGKLNKKKALKEEANKKTSLSQAEKCHHLNK-- 1019
OY 75 MLCSTLKQMKFLBORODETQKAREEAHRLKCKMK-----TMEQIELLLQSORSEVEEMI 129
DB 1020 -LKAKEPDQALDELBDNLEREKKVGDEKAKRKVEQDLKSQENVEDLERVKR-ELEENV 1077
OY 130 R-----DMGVQSAVEQLAVYCVSLKKEYEENIKERKATGELADRLKLDVSSR 178
DB 1078 RRKEAEISSLSKLEDSQNLVSQLRKIKELQARIEELELEELAERNAKAVEQKQRELN 1137
OY 179 SKLTLTNLTDLQA-----KLELSAQKDLQSA-----DOETISLRKSSD 218
DB 1138 RELEBERLDEDAAGATSAQIELELKKREBELKTRDLEEASLDQHEADQISLKRKHOD 1195

RESULT 11
TANA_XENLA
ID TANA_XENLA STANDARD; PRT; 1744 AA.
AC Q01550;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tanabrin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
```

```
OX NCBI_TaxID=8355;
RN {}
RP SEQUENCE FROM N.A.
RC TISSUE=Tadpole head;
RX MEDLINE=92398961; Pubmed=1524825;
RA Hemmati-Brivanlou A., Mann R.W., Harland R.M.;
RT "A protein expressed in the growth cones of embryonic vertebrate
neurons defines a new class of intermediate filament protein.";
RL Neuron 9:417-428(1992).
CC -1- TISSUE SPECIFICITY: GROWTH CONES OF EMBRYONIC VERTEBRATE NEURONS.
CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED IN THE NEURULA AND PERSIST
CC DURING EMBRYOGENESIS IN THE BRAIN, CRANIAL NERVES, AND SPINAL
CC CORD.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M99387; AAA49966.1; -.
DR PIR: JH0720; JH0720.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 2.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neutrone.
FT DOMAIN 1 12
FT DOMAIN 13 314
FT DOMAIN 13 314
FT DOMAIN 315 1744
FT DOMAIN 8 48
FT DOMAIN 49 60
FT DOMAIN 61 156
FT DOMAIN 157 179
FT DOMAIN 180 193
FT DOMAIN 194 199
FT DOMAIN 200 314
SQ SEQUENCE 1744 AA; 199561 MW; 6502EAC9F6C4E93 CRC64;
```

```
Query Match 14.9%; Score 160.5; DB 1; Length 1744;
Best Local Similarity 24.3%; Pred. No. 0.07;
Matches 57; Conservative 58; Mismatches 71; Indels 49; Gaps 8;
```

```
OY 12 AQEEENVLDAEFLKNELDVSKAQLSQKREKRSQALIIDRLDTEERNATVESLQNALN 71
DB 33 ALREEN-----ELLRKEIHSLSRSSKSEKRCWKKHHEEM-KLRDALDDGH-----REMY 80
OY 72 KAEMLCSTLKQMKFLBORODETQKAREEAHRLKCKMKTWQIELLLQSOR----- 122
DB 81 QAENVRDSIYEIELEFVQRCLEEKQAREDA-----KRELSKSKLLEETRAQIWLKERL 135
OY 123 -----SEVEEMIRDMGVQSAVEQLAVYCVSLKKEYEENIKERKATGELADRLKLDVSSR 169
DB 136 GQLEAELEDLIRHDEEKALMEE-----ELASFSQLENFRAPAVAFKVEVDVYARKLSEI 192
OY 170 LKLDVSSRSKLTNLTNLELQAKLELSAQK-----LQSAQDETISLRKSSD 217
DB 193 WQGAVEEYKSESVLEAGLSSESKENLRKVLLENKQNRLLQSLDKELVSLKMKRE 247

RESULT 12
MYSB_CAEEL
ID MYSB_CAEEL STANDARD; PRT; 1966 AA.
AC P02566;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain B (MHC B).
DE Myosin heavy chain B (MHC B).
GN UNC-54 OR MYO-4.
OS Caenorhabditis elegans.
```


OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83273600; PubMed=6576334;
 RA Karn J., Brenner S., Barnett L.;
 RT "Protein structural domains in the *Caenorhabditis elegans* unc-54
 RT myosin heavy chain gene are not separated by introns.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
 RN [2]
 RP SEQUENCE OF 850-1966 FROM N.A.
 RX MEDLINE=82272395; PubMed=7202124;
 RA McLachlan A.D., Karn J.;
 RT "periodic charge distributions in the myosin rod amino acid sequence
 RT match cross-bridge spacings in muscle.";
 RL Nature 299:226-231(1982).
 RN [3]
 RP SEQUENCE OF 1876-1966 FROM N.A.
 RX MEDLINE=83232892; PubMed=6571695;
 RA Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S.,
 RA Waterston R.H.;
 RT "The genes sup-7 x and sup-5 III of *C. elegans* suppress amber
 RT nonsense mutations via altered transfer RNA.";
 RL Cell 33:575-583(1983).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (LMC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
 CC C.ELEGANS.
 CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
 CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J01050; AAA28124.1; -;
 DR EMBL: V01494; CAA24738.1; -;
 DR PIR: A02992; MKRW.
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head. 1.
 DR Pfam: PF02736; Myosin_N. 1.
 DR Pfam: PF01576; Myosin_tail. 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head. 1.
 DR SMART: SM00242; MYSC. 1.
 KM APP-binding: Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM APP-binding: Methylation; Alkylation; Multigene family.
 FT DOMAIN 1 850 MYOSIN HEAD-LIKE.
 FT DOMAIN 851 1966 COILED COIL (POTENTIAL).
 FT DOMAIN 851 1164 ALPHA-HELICAL TAILPIECE (S2).
 FT DOMAIN 1165 1176 HINGE.
 FT DOMAIN 1165 1966 LIGHT MEROMYOSIN (LMW).

FT NP_BIND 177 184 ATP (BY SIMILARITY).
 FT DOMAIN 665 687 ACTIN-BINDING.
 FT DOMAIN 769 783 ACTIN-BINDING.
 FT MOD_RES 128 128 METHYLATION (TRI-) (POTENTIAL).
 FT MOD_RES 705 705 MOD_RES
 FT MOD_RES 715 715 ALKYLATION (SH-1).
 FT CONFLICT 1337 1337 ALKYLATION (SH-2).
 FT CONFLICT 1880 1880 E -> R (IN REF. 2).
 FT CONFLICT 1880 1880 I -> L (IN REF. 2).
 SQ SEQUENCE 1966 AA; 225125 MW; B66F0BB2FE27B67F CMC64;
 Query Match 14.9%; Score 160.5; DB 1; Length 1966;
 Best Local Similarity 25.1%; Pred. No. 0.079;
 Matches 59; Conservative 41; Mismatches 78; Indels 57; Gaps 9;
 QY 3 IINKLFDLAOEENYVDAEFLKNELDVKAQISQKREKRDSQALITDRTDLEEFENAT 62
 DB 1014 INRKLMEDLOSSEEDK-----GNHQNKKVAKLEQ-----TLDDLEDSLREKRA 1056
 QY 63 VESIQNMLNKAEMCSFLKQMKFLEORDET-KQAREEAHRLCKMKTKTEQIELLLSQ 121
 DB 1057 RADDKQKRRVE-----GELKIAQENIDSGQRHDLNLLKKESELHVSRRLEDE 1109
 QY 122 RSEYEEMIRDMGYQGSAAVEDLAVYCVSLKREYENLKEARRATGELADRLKDIYSSRSKL 181
 DB 1110 QALVSKIQRIQKQDSRISL-----LEELLENERSRSK-----ADRAKSDL---QREL 1155
 QY 182 KTLNTELDQ-----AKLELRSAQKDLQSA---DOETLSLRKSSDD 218
 DB 1156 EELGEKLEDEGGATAAGAAVEYNNKREELAKLRDLLEENAMNHNENQGLGRLKRRHTD 1210
 RESULT 13
 MYHB_MOUSE STANDARD; PRT; 1972 AA.
 ID MYHB_MOUSE
 AC 008638; 008639; 062462; 064195;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, smooth muscle isoform (SMWHC).
 GN MYH11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Uterus;
 RX MEDLINE=97242182; PubMed=9125171;
 RA Hasegawa K., Arakawa E., Oda S., Matsuda Y.;
 RT "Molecular cloning and expression of murine smooth muscle myosin heavy
 RT chains.";
 RL Biochem. Biophys. Res. Commun. 232:313-316(1997).
 RN [2]
 RP SEQUENCE OF 1-368 FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=95008063; PubMed=7923625;
 RA Milano J.M., Cserjesi P., Liqon K.L., Perissamy M., Olson E.N.;
 RT "Smooth muscle myosin heavy chain exclusively marks the smooth muscle
 RT lineage during mouse embryogenesis.";
 RL Circ. Res. 75:803-812(1994).
 RN [3]
 RP SEQUENCE OF 1-126 FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=96172919; PubMed=8593698;
 RA Suzuki T., Kim H.S., Kurabayashi M., Hamada H., Fujii H., Aikawa M.,
 RA Metanabe M., Watanabe N., Sakomura Y., Yazaki Y., Nagai R.;
 RT "Preferential differentiation of pig mouse embryonal carcinoma cells
 RT into smooth muscle cells. Use of retinoic acid and antisense against
 RT the central nervous system-specific POU transcription factor Brn-2.";
 RL Circ. Res. 78:395-404(1996).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2

```

CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2: ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PPM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC AKTYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHADED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D85923; BAA19690.1; -
DR EMBL; D85924; BAA19691.1; -
DR EMBL; L25860; AAA67552.1; -
DR EMBL; S81516; AAB36168.1; -
DR HSSP; P10587; 1BR1.
DR MGD; MGI:102643; MYH11.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family; Alternative splicing.
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT DOMAIN 786 815 IQ.
FT DOMAIN 844 1934 COILED COIL (POTENTIAL).
FT DOMAIN 1935 1972 CARBOXYL-TERMINAL.
FT NP_BIND 178 185 ATP (POTENTIAL).
FT DOMAIN 661 683 ACTIN-BINDING (BY SIMILARITY).
FT DOMAIN 762 776 ACTIN-BINDING (BY SIMILARITY).
FT MOD_RES 129 129 METHYLATION (SH-1) (POTENTIAL).
FT MOD_RES 701 701 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 711 711 ALKYLATION (SH-2) (POTENTIAL).
FT VARSPPLIC 1930 1972 RGNASFPVSRAGRGVRENTGSEEMDARSPDNGTGA
SE -> GPPQETSG (IN ISOFORM 2).
FT CONFLICT 126 126 N -> D (IN REF. 3).
FT CONFLICT 161 161 A -> V (IN REF. 2).
FT CONFLICT 189 189 Q -> K (IN REF. 2).
SQ SEQUENCE 1972 AA; 227026 MW; A1398B3F5B1F15A CRC64;
Query Match 14.98; Score 160; DB 1; Length 1972;
Best Local Similarity 22.38; Pred. No. 0.084;
Matches 53; Conservative 53; Mismatches 91; Indels 48; Gaps 7;
OY 10 DLAAEEENVLDAAFLKNELDSVKRAQLSOK-----DREKRDQAIIITDLRLTEE 58
DB 1688 DLAMQLEDLAAAFARAKQADLEKEELAEELASSGRTNLTODEKRRRLAEARIAQLLEEL 1747

```

```

OY 59 RNATVESIQNALNKAEMLCSTLKKOMKFLERODETKOAREEAHR----- 103
DB 1748 EOGNEMAMSDVRKATLOAEQLSNELATERSTAQKNESARQOLEROKKELRSKLQVEGA 1807
OY 104 LKCMK-TRMEDILOLOSREVEEMIRDMKVGSAVQOLAVVLSLKEVENLKEARKA 162
DB 1808 VKARKLSTVVALLEAKIKQLEQVQEARER---QAAKSLKQKDKKLEVLLOVEDERK- 1863
OY 163 TGEIADRLLKDVSSRSKLTKLITNELQAKLE-----TRSAQKDSADQ-----E 208
DB 1864 ---MAEQYKDAEKGNGKRVKQKQLEAEHEESOCIANNRKRLQREIDEATESNEAGRE 1920
OY 209 ITSLLRK 215
DB 1921 VNALKSK 1927
RESULT 14
KINH_HUMAN STANDARD; PRT; 963 AA.
AC P33176;
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-1993 (Rel. 40, Last sequence update)
DE Kinesin heavy chain (ubiquitous kinesin heavy chain) (UKHC).
GN KIF5B OR KNS1 OR KNS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92296683; PubMed=1607388;
RA Navone F., Nicolas J., Hom-Booher N., Sparks L., Bernstein H.D.,
RT "Cloning and expression of a human kinesin heavy chain gene:
RT interaction of the COOH-terminal domain with cytoplasmic microtubules
RT in transfected CV-1 cells."
RL J. Cell Biol. 117:1263-1275(1992).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1-349.
RX MEDLINE=96195066; PubMed=8606779;
RA Kull F.J., Sablin E.P., Lau R., Fletcher R.J., Vale R.D.;
RT "Crystal structure of the kinesin motor domain reveals a structural
RT similarity to myosin."
RL Nature 380:550-555(1996).
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=94242426; PubMed=7514426;
RA Nicolas J., Navone F., Hom-Booher N., Vale R.D.;
RT "Cloning and localization of a conventional kinesin motor expressed
RT exclusively in neurons."
RL Neuron 12:1059-1072(1994).
CC -1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.
CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT
CC CHAINS.
CC -1- SUBCELLULAR LOCATION: UNIFORMLY DISTRIBUTED BETWEEN THE CELL BODY
CC AND THE PROCESSES IN THE NEURONS.
CC -1- TISSUE SPECIFICITY: FOUND IN NEWBORN AND ADULT BRAIN, LIVER,
CC KIDNEY, SPLEEN, HEART, LUNG AND SCIATIC NERVE.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-
CC TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF
CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
CC DIMERIZATION, AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
CC VESICLES AND MEMBRANOUS ORGANELLES.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
 FT DOMAIN 786 815 IO.
 FT DOMAIN 844 1934 COILED COIL (POTENTIAL).
 FT DOMAIN 1935 1972 CARBOXYL-TERMINAL.
 FT NR_BIND 178 185 ATP (POTENTIAL).
 FT DOMAIN 661 683 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 762 776 ACTIN-BINDING (BY SIMILARITY).
 FT MOD_RES 129 129 METHYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 701 701 ALKYLATION (SH-2) (POTENTIAL).
 FT MOD_RES 711 711 ALKYLATION (SH-2) (POTENTIAL).
 FT CONFLICT 887 889 EER -> NSE (IN REF. 3).
 FT CONFLICT 1263 1266 ELOS -> TISE (IN REF. 2).
 FT CONFLICT 1558 1558 T -> S (IN REF. 3).
 FT CONFLICT 1610 1611 KO -> NE (IN REF. 3).
 FT CONFLICT 1786 1786 A -> S (IN REF. 4).
 FT CONFLICT 1958 1958 T -> L (IN REF. 3).
 SQ SEQUENCE 1972 AA; 227338 MW; 67655BB2ACE1277 CRC64;

Query Match 14.8%; Score 159; DB 1; Length 1972;
 Best Local Similarity 23.9%; Pred. No. 0.095;
 Matches 59; Conservative 52; Mismatches 88; Indels 48; Gaps 9;

QY 10 DLAGEEENVDAEFLKNEELSVKQLSQK-----DREKDSQAIIIDTLRDTLEE 58
 DB 1688 DLMQJEDDLAAERARQADLEKELEBELASSLSGRNALODEKRRLREARIQAELEEE 1747
 QY 59 RNATVESLQNALNK-----AEMLC-----STLKKOMKFLBORDETKQAREEHAHL--- 104
 DB 1748 EOGNMEASDVVRKATQOAEOLSNELATERSTQAKNESARQOLEKONELSKLHEMGA 1807
 QY 105 -KCKMK-TMEQIELLLQSRESEVEMTRDMGVGSAVEQLAVVCSLKEKEYENKEARKA 162
 DB 1808 VKSKFKSTIALLEAKIQLEQVEQEARKE---QAATKSLKQKKKLKEILLQYDEDERK- 1863
 QY 163 TGEIADRLKKDLSRSKLTLTNLTLDQAKLE-----LRSQAKDLOSADQ-----E 208
 DB 1864 ---NAEOYKEQAEKGNKAVKQKLEBEAEESORINANRRKRLQREIDEATESNEAMGRE 1920
 QY 209 ITSLSRK 215
 DB 1921 VNALKSK 1927
 RESULT 16
 MYHB_RABIT
 ID MYHB_RABIT STANDARD; PRT: 1972 AA.
 AC P35748;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, smooth muscle isoform (SMHMC).
 GN MYH11.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92073350; Pubmed=1961735;
 RA Babilj P., Kelly C., Perlasamy M.;
 RT "Characterization of a mammalian smooth muscle myosin heavy-chain
 RT gene: complete nucleotide and protein coding sequence and analysis of
 RT the 5' end of the gene."
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10676-10680(1991).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (LHC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MHC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M77812; AAA31395.1; -.
 DR PIR: A41604; A41604.
 DR HSSP: P08799; IMMD.
 DR InterPro: IPR000048; IO.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_Tail.
 DR InterPro: IPR002017; Spectrin.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00663; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_Tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PSS0096; IQ; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW Multigene family.
 FT DOMAIN 1 784 MYOSIN HEAD-LIKE.
 FT DOMAIN 785 807 IO.
 FT DOMAIN 844 1934 COILED COIL (POTENTIAL).
 FT DOMAIN 1935 1972 CARBOXYL-TERMINAL.
 FT NR_BIND 178 185 ATP (POTENTIAL).
 FT DOMAIN 661 683 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 763 777 ACTIN-BINDING (BY SIMILARITY).
 FT MOD_RES 129 129 METHYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 701 701 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 711 711 ALKYLATION (SH-2) (POTENTIAL).
 SQ SEQUENCE 1972 AA; 227318 MW; 2061A22428BD6A4C CRC64;

Query Match 14.8%; Score 159; DB 1; Length 1972;
 Best Local Similarity 22.7%; Pred. No. 0.095;
 Matches 56; Conservative 52; Mismatches 91; Indels 48; Gaps 7;

QY 10 DLAGEEENVDAEFLKNEELSVKQLSQK-----DREKDSQAIIIDTLRDTLEE 58
 DB 1688 DLMQJEDDLAAERARQADLEKELEBELASSLSGRNALODEKRRLREARIQAELEEE 1747
 QY 59 RNATVESLQNALNK-----AEMLCSTLKKOMKFLBORDETKQAREEHAHLKCKMKTVE-- 112
 DB 1748 EOGNMEASDVVRKATQOAEOLSNELATERSTQAKNESARQOLEKONELSKLQEMGA 1807
 QY 113 -----QIELLOSQSEVEEMTRDMGVGSAVEQLAVVCSLKEKEYENKEARKA 162
 DB 1808 VKSKFKSTIALLEAKIQLEQVEQEARKE---QAAKALKQKDKKLKEMILQVEDEERK- 1863
 QY 163 TGEIADRLKKDLSRSKLTLTNLTLDQAKLE-----LRSQAKDLOSADQ-----E 208
 DB 1864 ---NAEOYKEQAEKGNKAVKQKLEBEAEESORINANRRKRLQREIDEATESNEAMGRE 1920
 QY 209 ITSLSRK 215
 DB 1921 VNALKSK 1927

```

RESULT 17
RASO_ARCFU STANDARD; PRT; 886 AA.
ID RASO_ARCFU
AC 029230;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA double-strand break repair rad50 ATPase.
GN RAD50 OR AF1032.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Richardson K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Fleischmann D.L., Kierlavage A.R., Graham D.E., Kyriides N.C.,
RA Fleischmann E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Kirkness S., Reich C.I., McNeil L.R., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sedow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC CC
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mre11 complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (by
CC similarity).
CC -1- SUBUNIT: Forms a complex with mre11 (by similarity).
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001032; MAB90211.1; -.
CC DR TIGR: AFI032; -.
CC DR InterPro: IPR003439; ABC_transport.
CC DR InterPro: IPR001238; RecF.
CC DR InterPro: IPR002017; Spectrin.
CC DR Pfam: PF00470; RecF. 1.
CC KM DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
CC FT NP_BIND 31 38 ATP (BY SIMILARITY).
CC FT DOMAIN 148 728 COILED COIL (POTENTIAL).
CC SQ SEQUENCE 886 AA; 103633 MW; D35641D499AAB58 CRC64;

```

Query Match 14.7%; Score 157.5; DB 1; Length 886;
 Best Local Similarity 22.3%; Pred. No. 0.051;
 Matches 56; Conservative 71; Mismatches 83; Indels 41; Gaps 9;

```

QY 1 KTIINKLFFDLQAEENVLDAEFLKNEIDSVKAO-----LSQKDRK-KRDSQAI 49
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 207 ESLEKESIEEVRNLSRLKEEHSRLSESRKQESSVLQVRLLEKRLREQLKEVY 266
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 50 DTLADTLERRAIVESLQNALNK---ADMCLSTLKKQMKFLQNR-----QDET 95
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 267 ERIED-LEKKAKEVLEKPKAERYSLLEKLSLSEINQALRDVEKRECDLTRRAGIQAK 325
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

QY 96 QAREARHLCKMKTMEQIE-----LLQSQRSEYEMIRMGVGSQAVEQLAVY 145
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 326 KAEDNSKLEETTRITRELELEREFKSHRLLETLKPKMDR---QGI-KAKLEENLT 381
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 146 CYSLEKEVEMLEKARKATGELADRLKKDLVSRSKLTMTLDELDAKLELRSAQKQSA 205
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 382 PDKEKMYDLSLKKAKEKEKTEKTKK-LIAKKSLSLTKRQAQLKKAIVELKSNRTPVC 440
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 206 DQETLSLRKS 216
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 441 GREIDEEHRKN 451
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 18
MYSN_ACACA STANDARD; PRT; 1509 AA.
ID MYSN_ACACA
AC P05659;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin II heavy chain, non muscle.
OS Acanthamoeba castellanii (Amoeba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBI_TaxID=5755;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87308395; PubMed=3040773;
RA Hammer J.A. III, Bowers B., Paterson B.M., Korn E.D.;
RT "Complete nucleotide sequence and deduced polypeptide sequence of a
RT nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT hinge in the rodlike tail.";
RL J. Cell Biol. 105:913-925(1987).
CC CC
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS APPRASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.
CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMATIC ACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y00624; CA68663.1; -.
CC DR PIR: A27224; A27224.
CC DR HSSP: P08799; IMND.
CC DR InterPro: IPR000048; IQ.
CC DR InterPro: IPR004009; Myosin_N.
CC DR InterPro: IPR001609; myosin_head.
CC DR Pfam: PF00612; IQ; 2.
CC DR Pfam: PF00063; myosin_head; 1.
CC DR Pfam: PF02736; Myosin_N; 1.
CC DR PRINTS: PR00193; MYOSINHEAVY.
CC DR ProDom: PD000355; myosin_head; 1.
CC DR SMART: SM00015; IQ; 1.
CC DR SMART: SM00242; MISC; 1.
CC DR PROSITE: PS50096; IQ; 1.
CC KM Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
CC Methylation; Alkylation; Phosphorylation; Multigene family.

```

```
FT DOMAIN 1 789 MYOSIN HEAD-LIKE.
FT DOMAIN 790 819 IQ.
FT DOMAIN 848 1509 COILED COIL (POTENTIAL).
FT DOMAIN 848 1226 ALPHA-HELICAL TAILPIECE (S2).
FT DOMAIN 1227 1252 HINGE.
FT DOMAIN 1253 1509 LIGHT MEROMYOSIN (LMM).
FT DOMAIN 1253 1482 ALPHA-HELICAL TAILPIECE (LMM).
FT DOMAIN 1483 1509 NONHELICAL TAILPIECE.
FT NP_BIND 182 189 ATP.
FT DOMAIN 660 682 ACTIN-BINDING.
FT DOMAIN 766 780 ACTIN-BINDING.
FT MOD_RES 133 133 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 700 700 AKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 1489 1489 PHOSPHORYLATION.
FT MOD_RES 1494 1494 PHOSPHORYLATION.
FT MOD_RES 1499 1499 PHOSPHORYLATION.
SO SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

Query Match 14.6%; Score 156.5; DB 1; Length 1509;
Best Local Similarity 20.9%; Pred. No. 0.099;
Matches 59; Conservative 59; Mismatches 91; Indels 73; Gaps 7;

QY 10 DLADAEENVLD-----AEFLKNEIDSV-----KAOLS 36
D 1017 DALDAENISETLRSKLTNTERGADVNEIDVYATKLEKTKKSLSEELATRAOLE 1076
QY 37 OKDEKRSOALIDTRNLEERNATVESLONALNKAEMLSTLKKOMKFLERODEKQ 96
D 1077 EEKSKKEAASKAKOLGQOLEANSEVDLSKSAEKSLSKTADONRDDEOLEDERT 1136
QY 97 AR---EEAHLKCKMKTMEQIELLSQ-----RSEVEEMRDGVGQ----- 136
D 1137 VRANVDKOKKLEKLETELEQVYALDQKNAAQAATLTQVDETRRLLEEAASAR 1196
QY 137 -----SAVEQLAYCVSLKEYEENLEKARKATGELADRLKKDLY-----SSRSKLT 183
D 1197 LEKERKNALDEVAQLTADLDERDSGAQORRKLNRRISELEENAPKGTGASSEYVK 1256
QY 184 LTELDOAKLELSAO-----KDLSADQETLSLRKSD 218
D 1257 LEGELERLEEBELTRQEARAAAEKNLKANLELELROEAD 1298

RESULT 19
VDP_HUMAN STANDARD; PRT; 962 AA.
ID VDP_HUMAN
AC 060763;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE General vesicular transport factor p115 (Transcytosis associated
protein) (TAP) (Vesicle docking protein).
GN VDP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mamalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. AND PHOSPHORYLATION SITE SER-942.
RX MEDLINE-98148093; PubMed-9478999;
RA Sonda M., Misumi Y., Yano A., Takami N., Ikehara Y.;
RT "Phosphorylation of the vesicle docking protein p115 regulates its
RT association with the Golgi membrane.";
RL J. Biol. Chem. 273:5385-5388(1998).
CC -1- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR
CC INTERSTITIAL FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO
CC THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY
CC INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR
CC AND TARGET MEMBRANES IN PROXIMITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES
CC BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE.
```

```
CC -1- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-
CC COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.
CC -1- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER;
CC PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.
CC DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE;
CC PHOSPHORYLATION PROMOTES DISSOCIATION.
CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D86326; BAA25300.1; -.
DR MIM; 603344; -.
DR InterPro; IPR00225; Armadillo.
DR PROSITE; PS50176; ARM_REPEAT; UNKNOWN_1.
KW Transport; Protein transport; Golgi stack; Membrane; Coiled coil;
KW Phosphorylation.
FT DOMAIN 1 637 GLOBULAR HEAD.
FT DOMAIN 638 930 COILED COIL (POTENTIAL).
FT DOMAIN 935 962 ASP/GLU-RICH (ACIDIC).
FT MOD_RES 942 942 PHOSPHORYLATION.
FT MOD_RES 942 942 S->A: LOSS OF PHOSPHORYLATION.
SO SEQUENCE 962 AA; 107906 MW; 2E748FC1BC2B942 CRC64;
```

```
Query Match 14.4%; Score 155; DB 1; Length 962;
Best Local Similarity 19.8%; Pred. No. 0.076;
Matches 64; Conservative 57; Mismatches 93; Indels 110; Gaps 6;
```

```
QY 3 IINKLFDLADAEENVLDLAEFLKNEIDSVKAOLOSKREKDSOAIIDTLDTLEERNAT 62
D 608 IFDEHFTLVLELGVYTKAYIKSSEEDKKEEVKKTLEQHDN--IYHYKNMIREDDIQ 665
D 63 VESLONALN-----KAEMLSCTLKMKOMKFLERODETKOAR----- 98
QY 666 LLELRQOVSTLKCONELOLQAVTQOVSOIQOHKQYINLTKLOLQKDNHOGSYSEGAOMN 725
D 99 ---EEAHLKCKMKTMEQIELLSQSEVEEMIRDMVGOSA----- 138
D 726 GIOPEEIGRLREIELEIKRNOELQLOLTKDSMIENKKSOTSOTGNEOSSAIVARDE 785
QY 139 -VEQLAYCVSLK----- 150
D 786 QVAELKQELATLKSQNSQVEYTRKLOTEKQELLQKTEAFKSVVOGETITATKTTD 845
QY 151 -----KEYENLEKARKATGELADRLKKDLYSSRSKLTLMTELDOAKLELSAOKD 201
D 846 VEGRLSALQETKEKKNKRLSEERRAKIQDSSNSTIALLQTEKDKELELTDSKKE 905
QY 202 -----LQSDAQETLSLRKSD 218
D 906 ODDLVLVLLADQDKILSLKNLKD 929

RESULT 20
MYS_A_DROME STANDARD; PRT; 2411 AA.
ID MYS_A_DROME
AC P05661;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, muscle.
GN MHC.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
```

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89384556; PubMed=2506434;
 RA George E.L., Ober M.B., Emerson C.P. Jr.;
 RT "Functional domains of the Drosophila melanogaster muscle myosin
 heavy-chain gene are encoded by alternatively spliced exons."
 RL Mol. Cell. Biol. 9:2957-2974(1989).
 RN [2]
 RP SEQUENCE OF 1-312 FROM N.A.
 RX MEDLINE=87280141; PubMed=3038896;
 RA Wassenberg D.R. II, Kronert W.A., O'Donnell P.T., Bernstein S.I.;
 RT "Analysis of the 5' end of the Drosophila muscle myosin heavy chain
 gene. Alternatively spliced transcripts initiate at a single site and
 intron locations are conserved compared to myosin genes of other
 organisms."
 RT J. Biol. Chem. 262:10741-10747(1987).
 RN [3]
 RP SEQUENCE OF 486-881 FROM N.A.
 RX STRAIN-CANTON-S; TISSUE=Embryonic muscle;
 RC MEDLINE=91330870; PubMed=1907912;
 RA Kronert W.A., Edwards K.A., Roche E.S., Wells L., Bernstein S.I.;
 RT "Muscle-specific accumulation of Drosophila myosin heavy chains: a
 splicing mutation in an alternative exon results in an isoform
 substitution."
 RT EMBO J. 10:2479-2488(1991).
 RL
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- ALTERNATIVE PRODUCTS: AS DROSOPHILA HAS A SINGLE MUSCLE MHC GENE,
 MHC ISOFORMS MUST RELY ON ALTERNATIVE RNA SPLICING THAT LEADS TO
 DIFFERENCES IN THE C-TERMINUS OF THE VARIOUS MHC PROTEINS.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M61229; AAA28686.1; ALT_SEQ.
 DR EMBL: M61229; AAA28687.1; ALT_SEQ.
 DR EMBL: J02788; AAA28706.1; ALT_SEQ.
 DR EMBL: J02788; AAA28707.1; ALT_SEQ.
 DR EMBL: X60196; CAA42752.1; ALT_SEQ.
 DR EMBL: X60196; CAA42753.1; ALT_SEQ.
 DR EMBL: X60196; CAA42754.1; ALT_SEQ.
 DR PIR: A28492; A28492.
 DR PIR: A32491; A32491.
 DR PIR: B32491; B32491.
 DR HSSP: P08799; 1MND.
 DR FlyBase: Fggn0002741; Mhc.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR002928; Myosin_N.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 2.
 DR Pfam: PF00663; myosin_head; 7.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSTINHEAVY.
 DR PRODOM: PD000355; myosin_head; 5.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.

KW Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM ATP-binding; Calmodulin-binding; Alternative splicing;
 KW Multigene family.
 FT DOMAIN 1 MYOSIN HEAD-LIKE.
 FT DOMAIN 1205 1232 IQ.
 FT DOMAIN 1233 2376 COILED COIL (POTENTIAL).
 FT NP_BIND 227 234 ATP (BY SIMILARITY).
 FT VARSPIC 2385 2385 P -> I (IN SHORT ISOFORM).
 FT VARSPIC 2386 2411 MISSING (IN SHORT ISOFORM).
 FT CONFLICT 43 44 EK -> RE (IN REF. 2).
 SQ SEQUENCE 2411 AA: 276434 MW: 804FEC1BB8D310A9 CRC64;
 Query Match 14.4%; Score 154.5; DB 1; Length 2411;
 Best Local Similarity 22.8%; Pred. No. 0.21; Indels 53; Gaps 9;
 Matches 59; Conservative 56; Mismatches 9;
 QY 10 DLAAQEEENVLD--AEFLKNELDGSKAQLSQKDEKRPDSQAIIPTLRDPLFEERNATVESIQ 67
 Db 2113 DDAREQGISERRANMLQNELESRTLLLEGADRGRAQEGELADAHQELNEVSAQNASIS 2172
 QY 68 NALNKAEMLCSTLKKQKFL----EQRDETKAREBAHFLKCKMTMEDIELLOSQRS 123
 Db 2173 AARKKLESELQTLHSDLELDLNEAKNSEKAKKAMDADAELRAEDPDHQAOTKRLK 2232
 QY 124 EVEEMTRDMGV-----GQSAVEQLAVYCVSLKKEYE-----NLRK 160
 Db 2233 ALDQIKELQVRLDEAEANLKGGRKATQKLEQRYLELDGEBORRNADAKNLKSE 2292
 QY 161 KATGELA----DLK---KDLVSS-RSKLKTLMLELDAQ-----LELSAQKDLQ 203
 Db 2293 RRVKELSFQSEEDSRKRNHERMQDLVYDKLQIKYKRIEBAEELALNLAKPKQAQELE 2352
 QY 204 S-----ADQETSLKK 215
 Db 2353 EAERADLAQALSKFRK 2371
 RESULT 21
 ID EP15_HUMAN STANDARD; PRT; 896 AA.
 AC P42566;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Epidermal growth factor receptor substrate 15 (Protein EP15) (AF-1p
 protein).
 GN EP15 OR AFIP.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=melanoma;
 RX MEDLINE=94239734; PubMed=8183552;
 RA Wong W.T., Kraus M.H., Carlomagno F., Zelano A., Druck T.,
 Croce C.M., Huebner K., di Fiore P.P.;
 RT "The human eps15 gene, encoding a tyrosine kinase substrate, is
 conserved in evolution and maps to 1p31-p32.";
 RL Oncogene 9:1591-1597(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94181254; PubMed=8134107;
 RA Bernard O.A., Mauchault M., Mecucci C., van den Berghe H.,
 Berger R.;
 RT "A novel gene, AF-1p, fused to HRX in t(1;11)(p32;q23), is not
 related to AF-4, AF-9 nor ENL.";
 RL Oncogene 9:1039-1045(1994).
 CC -1- FUNCTION: INVOLVED IN CELL GROWTH REGULATION. MAY PLAY A ROLE IN
 SIGNAL TRANSDUCTION AND MITOGENICITY.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED.


```

CC -1- P-TM: PHOSPHORYLATED ON TYROSINE BY EGFR.
CC -1- DISEASE: INVOLVED IN A T(1;11)(P32;Q23) CHROMOSOMAL TRANSLOCATION
CC IN ACUTE LEUKEMIAS CAUSING FUSION TO THE TRITHORAX (MLL OR HRX)
CC GENE PRODUCT WHICH CONTAINS DNA-BINDING MOTIFS RESULTING IN A
CC ROGUE ACTIVATOR PROTEIN.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U07707; AAA52101.1; -.
DR EMBL: Z29064; CAA82305.1; -.
DR HSSP: P02633; 1BOD.
DR MIM: 600051; -.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000261; EF-hand.
DR InterPro: IPR003903; UIM.
DR Pfam: PF00036; ehand; 4.
DR SMART: SM00054; EFN; 3.
DR SMART: SM00027; EH; 3.
DR PROSITE: PS00018; EF_HAND; 2.
KW Phosphorylation; Calcium-binding; Repeat; Chromosomal translocation;
KW Proto-oncogene.
FT CA_BIND 173 184 EF_HAND 1 (POTENTIAL).
FT CA_BIND 236 247 EF_HAND 2 (POTENTIAL).
FT DOMAIN 599 827 15 X 3 AA REPEATS OF D-P-F.
FT REPEAT 599 601 1.
FT REPEAT 623 625 1.
FT REPEAT 629 631 3.
FT REPEAT 634 636 4.
FT REPEAT 640 642 5.
FT REPEAT 645 647 6.
FT REPEAT 651 653 7.
FT REPEAT 664 666 8.
FT REPEAT 672 674 9.
FT REPEAT 692 694 10.
FT REPEAT 709 711 11.
FT REPEAT 737 739 12.
FT REPEAT 798 800 13.
FT REPEAT 804 806 14.
FT REPEAT 825 827 15.
FT DOMAIN 768 850 PRO-RICH.
FT MOD_RES 132 132 PHOSPHORYLATION (BY TYR-KINASES)
FT CONFLICT 822 822 M -> I (IN REF. 2).
FT SEQUENCE 896 AA; 98673 MW; A1B9PB04A07FABEB CRC64;

```

Query Match 14.3%; Score 154; DB 1; Length 896;
 Best Local Similarity 22.6%; Pred. No. 0.08;
 Matches 53; Conservative 42; Mismatches 60; Indels 80; Gaps 7;

```

RESULT 22
MYHA_HUMAN STANDARD; PRT; 1976 AA.
AC P35580;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, nonmuscle type B (cellular myosin heavy chain,
DE type B) (Nonmuscle myosin heavy chain-B) (NMHC-B).
GN MYH10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96025307; Pubmed=7499478;
RA Phillips C.L., Yamakawa K., Adelstein R.S.;
RT "Cloning of the cDNA encoding human nonmuscle myosin heavy chain-B and
RT analysis of human tissues with isoform-specific antibodies."
RL J. Muscle Res. Cell Motil. 16:379-389(1995).
RN [2]
RP SEQUENCE OF 63-722 FROM N.A.
RX MEDLINE=91316803; Pubmed=1860190;
RA Simons M., Wang M., McBride O.W., Kawamoto S., Yamakawa K.,
RA Gula D., Adelstein R.S., Weir L.;
RT "Human nonmuscle myosin heavy chains are encoded by two genes located
RT on different chromosomes."
RL Circ. Res. 69:530-539(1991).
CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
CC CAPING.
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M69181; AAA99177.1; -.
DR PIR: B61231; B61231.
DR HSSP: P08799; 1LVK.
DR MIM: 160776; -.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR000409; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Alkylation; Multigene family.
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT DOMAIN 786 815 IQ.

```

FT DOMAIN 845 1976 COILED COIL (POTENTIAL).
 FT NE_BIND 178 185 ATP (POTENTIAL).
 FT MOD_RES 701 701 ALKYLIATION (SH-1) (POTENTIAL).
 FT MOD_RES 711 711 ALKYLIATION (SH-2) (POTENTIAL).
 SQ SEQUENCE 1976 AA, 228938 MW, B2BB87F35EAL24F CRC64;

Query Match 14.3%; Score 154; DB 1; Length 1976;
 Best Local Similarity 21.0%; Pred. No. 0.18;
 Matches 64; Conservative 58; Mismatches 93; Indels 90; Gaps 10;

QY 4 INKLFDLAEEENV-----LDAEF-LKNEIDSVKAQDSQK 38
 DB 1078 IDEIKLQALAKKEELQALARGDEDETLKNNALKVRELQAOIQLQEDFESEKASRNKA 1137
 QY 39 DREKRSQAIIDRLDLEERNATVESLQNALNKAEMLCSTLKK-----QMKFL 87
 DB 1138 EKQKRDLSSEELAKTELEDTLDTTAAQQLRTKREQVVALKLEETKNNHQAQIDM 1197
 QY 88 EORQ----DETKQAREEAHRLK-----CKMKTMEQIELLLQSQR--- 122
 DB 1198 RQRNATALEELSEQLQAKKFKFKALEKNKQGLTDNKEACEVAVLQOVAESHKRKKL 1257
 QY 123 -SEVEEIRDMGVQSAVEQLAVYCVSLKKEYEN-----LKEARKA----- 162
 DB 1258 DAQVOELHAYVEGDRRLRVELAEKASKLQNELDNVSTLLEAEKKGKIFAKDAASLESQ 1317
 QY 163 --GGLADRLKLDVSSRSKLTNTNT-----LDQAKLELRSGKQKLO---SADDEITSLR 213
 DB 1318 QDTQELQOETTRKLNLSRIROLEEKNLSLQEQEEEEKARKNLEKOVIALQSOLADTK 1377
 QY 214 KKSDD 218
 DB 1378 KKVD 1382

RESULT 23
 MYSID_CAEEL STANDARD; PRT; 1938 AA.
 ID MYSID_CAEEL
 AC P02567; 019674;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain D (MHC D).
 GN MYO-1 OR R06C7.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidea;
 OC Rhabdilitidae; Peloderiinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=89178677; PubMed=2926820;
 RA Dlibb N.J., Maruyama I.N., Krause M., Karn J.;
 RT "Sequence analysis of the complete Caenorhabditis elegans myosin
 heavy chain gene family.";
 RL J. Mol. Biol. 205:603-613(1989).
 RN [2]
 RP SEQUENCE OF 34-1795 FROM N.A.
 RX MEDLINE=83273600; PubMed=6576334;
 RA Karn J., Brenner S., Barnett L.;
 RT "Protein structural domains in the Caenorhabditis elegans unc-54
 myosin heavy chain gene are not separated by introns.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
 RN [3]
 RP SEQUENCE OF 115-365 AND 1492-1763 FROM N.A.
 RX MEDLINE=85201409; PubMed=3888374;
 RA Karn J., Dlibb N.J., Miller D.M.;
 RT "Cloning nematode myosin genes.";
 RL Cell Muscle Motil. 6:185-237(1985).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;

RA Gardner A., McMurray A.;
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (LMC)
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MHC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE PHARYNGEAL MUSCLE.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPE
 SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
 C.ELEGANS.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X08065; CAA30854.1; -;
 DR EMBL; M37232; AAA28119.1; -;
 DR EMBL; M37234; AAA28120.1; -;
 DR EMBL; Z71266; CAA95848.1; -;
 DR EMBL; Z71261; CAA95848.1; JOINED.
 DR EMBL; Z71261; CAA95806.1; -;
 DR EMBL; Z71266; CAA95806.1; JOINED.
 DR PIR; S02772; MMKWL.
 DR HSSP; P08799; 1MND.
 DR WormBep; R06C7.10; CE06253.
 DR InterPro; IPR004009; Myosin_N.
 DR InterPro; IPR002928; Myosin_tail.
 DR InterPro; IPR001609; myosin_head.
 DR Pfam; PF00063; myosin_head; 1.
 DR Pfam; PF02736; Myosin_N; 1.
 DR Pfam; PF01576; Myosin_tail; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR ProDom; PD000355; myosin_head; 1.
 DR SMART; SM00242; MYSC; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM ATP-binding; Methylation; Alkylation; Multigene family.
 FT DOMAIN 1 845
 FT DOMAIN 846 1938
 FT DOMAIN 846 1170
 FT DOMAIN 1171 1938
 FT DOMAIN 846 1938
 FT NP_BIND 177 184
 FT DOMAIN 660 682
 FT DOMAIN 764 778
 FT MOD_RES 128 128
 FT MOD_RES 700 700
 FT MOD_RES 710 710
 FT CONFLICT 94 94
 FT CONFLICT 98 98
 FT CONFLICT 377 377
 FT CONFLICT 389 390
 FT CONFLICT 391 391
 FT CONFLICT 408 408
 FT CONFLICT 474 474
 FT CONFLICT 577 577
 FT CONFLICT 681 681
 FT CONFLICT 1373 1373
 FT CONFLICT 1659 1659
 SQ SEQUENCE 1938 AA, 223255 MW, 387399C8F63A4CFA CRC64;
 V -> D (IN REF. 2).
 V -> D (IN REF. 4).
 V -> D (IN REF. 4).
 W -> N (IN REF. 2).
 Q -> G (IN REF. 2).
 L -> F (IN REF. 4).
 I -> N (IN REF. 4).
 S -> D (IN REF. 2).
 E -> Q (IN REF. 3).
 DV -> GD (IN REF. 2).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 RODLIKE TAIL (S2 AND LHM DOMAINS).
 ALPHA-HELICAL TAILPECE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLATION (SH-1).
 ALKYLATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4

DR InterPro: IPR000048; IQ.
 DR InterPro: IPR000409; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF006612; IQ_2.
 DR Pfam: PF000663; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW Multigene family.
 FT DOMAIN 1 784 MYOSIN HEAD-LIKE.
 FT DOMAIN 785 814 IQ.
 FT DOMAIN 843 1938 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP (POTENTIAL).
 FT DOMAIN 659 681 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 761 775 ACTIN-BINDING (BY SIMILARITY).
 FT MOD_RES 130 130 METHYLATION (TR1-) (POTENTIAL).
 FT MOD_RES 699 699 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).
 SQ SEQUENCE 1938 AA; 223678 MW; 1F6D006416381CDS CRC64;

Query Match 14.28; Score 153; DB 1; Length 1938;
 Best Local Similarity 24.9%; Pred. No. 0.2;
 Matches 59; Conservative 40; Mismatches 74; Indels 64; Gaps 9;

OY 20 DAERLKNEISVKAOLSOQREKRSQAIIID---TLRDLTEENNAVESIQNALNKAEM 75
 DB 1382 DAIO RTELEBAKKKLRQAEAEKETETANSKCSLEKTKRQLOGEVEDLMDLERSHT 1441
 OY 76 LCSTL-KKQKF-----LEQRDETKQAREAHRLCKM-----KTMEQIEL 116
 DB 1442 ACATLDKKRNFVKVLAEMWKQKDESOALELAQKESRSLSTELFKRNVAEEVVDQLET 1501
 OY 117 L-----LQSORSVEEMINDMGVGSQSAVEQLAVYCVSLKKEYENLKAARKAGELADRL 170
 DB 1502 LRRENKMOEISDLTEQIETG-----KNQDAEK-TKKLYEOE 1540
 OY 171 KKD-----LVSSRSKLTNTLDOAKLELSAQKDLQASQDITSLRKS 216
 DB 1541 KSDLQALVEEVGSLHEBSKILRVQLSEL--DKVTKEDELEQLKRN 1594
 RESULT 26
 MYSA CAEL STANDARD; PRT; 1969 AA.
 ID MYSA CAEL
 AC P12844;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Myosin heavy chain A (MHC A).
 GN MYO-3
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=89118677; PubMed=2926820;
 RA Dibb N.J., Maruyama I.N., Krause M., Karn J.;
 RT "Sequence analysis of the complete Caenorhabditis elegans myosin
 heavy chain gene family.";
 RL J. Mol. Biol. 205:603-613(1989).
 CC -!- FUNCTION: MUSCLE CONTRACTION.
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)

CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -!- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
 CC C.ELEGANS.
 CC -!- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
 CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X08067; CA30856.1; -.
 DR PTR: S02771; S02771.
 DR HSP: P08799; ILVK.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR000409; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF006612; IQ; 1.
 DR Pfam: PF000663; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW APP-binding; Methylation; Alkylation; Multigene family.
 FT DOMAIN 1 793 MYOSIN HEAD-LIKE.
 FT DOMAIN 794 823 IQ.
 FT DOMAIN 857 1969 COILED COIL (POTENTIAL).
 FT NP_BIND 179 186 ATP.
 FT DOMAIN 667 689 ACTIN-BINDING.
 FT DOMAIN 770 784 ACTIN-BINDING.
 FT MOD_RES 130 130 METHYLATION (TR1-) (POTENTIAL).
 FT MOD_RES 707 707 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 717 717 ALKYLATION (SH-2) (POTENTIAL).
 SQ SEQUENCE 1969 AA; 225509 MW; 64577BBAF7EAD80A CRC64;

Query Match 14.18; Score 152; DB 1; Length 1969;
 Best Local Similarity 24.38; Pred. No. 0.23;
 Matches 62; Conservative 47; Mismatches 98; Indels 48; Gaps 8;

OY 5 NKLFEFLAEEENVLDAEFLKNEISVKAOLSOQ-----DREKRS 45
 DB 898 NALFLSLTEKANLADAEERNEKLNQKATLESLSITQGLEDMQERNEDLAQKKKTD 957
 OY 46 QAITDT-----LRDTLE---RNATVESIQNALNKAEMLCSTLKKOMFLEGRD 92
 DB 958 QELSDTKKHVQDELSTLRKAEQEKSRDHNIRLQDEMANQDEAVAKLNKKK---HQEE 1014
 OY 93 ETQAREEAHRLCKMKTMEQIELLLQSORSVEEM-----RDMGVGSQSAVEQLAVYCVS 148
 DB 1015 SNRKLNDLQSEEDKVNHLKTRNKLEQOMDELENDREKRSKGLTEKKRVE---GD 1071
 OY 149 LKKEYENLKAARKAGELADRLKK---DLVSSRSKLTNTLNT---ELDQAKLELSAQKDL 202

Db 1072 LKVAENIDEITKQKHVETLTKREEDLHHTNKLAKLAENSLIAKLQRLIELFARNNEL 1131
 QY 203 QSAQOETSLEKSD 217
 Db 1132 EEELEARNRSKSD 1146

RESULT 27
 H1P1_HUMAN

ID H1P1_HUMAN STANDARD; PRT; 995 AA.
 AC 000291: 000328;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Huntingtin interacting protein 1 (H1P-1) (Fragment).
 GN H1P1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hug A.H.M., Nichol K., Osborne L., Scherer S.W., Squitieri F.,
 RA Hayden M.R.;
 RT "genomic organization of the human H1P1 gene and its exclusion as a
 RT candidate gene in a family diagnosed with Huntington disease without
 RT CAG expansion." to the EMBL/GenBank/DBJ databases.
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 82-995 FROM N.A.
 RX MEDLINE=97285121; PubMed=9140394;
 RA Kishiyama K., Kazem-Esfarjani P., Lynn F.C., Wellington C.,
 RA Metzler M., Goldberg Y.P., Kanazawa I., Geitz R.D., Hayden M.R.;
 RT "H1P1, a human homologue of S. cerevisiae Sla2p, interacts with
 RT membrane-associated huntingtin in the brain."
 RL Nat. Genet. 16:44-53(1997).
 RN [3]
 RP SEQUENCE OF 82-159 FROM N.A.
 RA Bradshaw H., Hinds K., Harrison M.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 203-602 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97227296; PubMed=9147654;
 RA Wanker E.E., Royra C., Scherzinger E., Hasenbank R., Waelter S.,
 RA Taft D., Colicelli J., Lehnach H.;
 RT "H1P-1: a huntingtin interacting protein isolated by the yeast two-
 RT hybrid system."
 RL Hum. Mol. Genet. 6:487-495(1997).
 RN [5]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=20515263; PubMed=11063258;
 RA Chopra V.S., Metzler M., Rasper D.M., Engqvist-Goldstein A.E.Y.,
 RA Singaraja R., Gan L., Fichter K.M., McCutcheon K., Drubin D.,
 RA Nicholson D.W., Hayden M.R.;
 RT "H1P2 is a non-proapoptotic member of a gene family including H1P1,
 RT an interacting protein with huntingtin."
 RL Mamm. Genome 11:1006-1015(2000).
 CC -1- FUNCTION: MAY PLAY A FUNCTIONAL ROLE IN THE CELL FILAMENT
 CC NETWORKS.
 CC -1- SUBUNIT: BINDS HUNTINGTIN. THIS INTERACTION IS RESTRICTED TO THE
 CC BRAIN.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. MEMBRANE-ASSOCIATED PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; H1P1-1 AND H1P1-2. ARE PRODUCED
 CC BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED WITH THE HIGHEST LEVEL
 CC IN BRAIN.
 CC -1- MISCELLANEOUS: THE AFFINITY OF THE HUNTINGTIN PROTEIN-H1P1
 CC INTERACTION IS INVERSELY CORRELATED TO THE LENGTH OF THE
 CC POLYGLUTAMINE TRACT ADDED TO THE HUNTINGTIN PROTEIN IN HUNTINGTON
 CC DISEASE.

CC -1- SIMILARITY: BELONGS TO THE SLA2 FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF052288; AAC33564.1; -
 DR EMBL; AF052261; AAC33564.1; JOINED.
 DR EMBL; AF052262; AAC33564.1; JOINED.
 DR EMBL; AF052263; AAC33564.1; JOINED.
 DR EMBL; AF052264; AAC33564.1; JOINED.
 DR EMBL; AF052265; AAC33564.1; JOINED.
 DR EMBL; AF052266; AAC33564.1; JOINED.
 DR EMBL; AF052267; AAC33564.1; JOINED.
 DR EMBL; AF052268; AAC33564.1; JOINED.
 DR EMBL; AF052269; AAC33564.1; JOINED.
 DR EMBL; AF052270; AAC33564.1; JOINED.
 DR EMBL; AF052271; AAC33564.1; JOINED.
 DR EMBL; AF052272; AAC33564.1; JOINED.
 DR EMBL; AF052273; AAC33564.1; JOINED.
 DR EMBL; AF052274; AAC33564.1; JOINED.
 DR EMBL; AF052275; AAC33564.1; JOINED.
 DR EMBL; AF052276; AAC33564.1; JOINED.
 DR EMBL; AF052277; AAC33564.1; JOINED.
 DR EMBL; AF052278; AAC33564.1; JOINED.
 DR EMBL; AF052279; AAC33564.1; JOINED.
 DR EMBL; AF052280; AAC33564.1; JOINED.
 DR EMBL; AF052281; AAC33564.1; JOINED.
 DR EMBL; AF052282; AAC33564.1; JOINED.
 DR EMBL; AF052283; AAC33564.1; JOINED.
 DR EMBL; AF052284; AAC33564.1; JOINED.
 DR EMBL; AF052285; AAC33564.1; JOINED.
 DR EMBL; AF052286; AAC33564.1; JOINED.
 DR EMBL; AF052287; AAC33564.1; JOINED.
 DR EMBL; AF052288; AAC33564.1; JOINED.
 DR EMBL; AF052289; AAC33564.1; JOINED.
 DR EMBL; AF052290; AAC33564.1; JOINED.
 DR EMBL; AF052291; AAC33564.1; JOINED.
 DR EMBL; AF052292; AAC33564.1; JOINED.
 DR EMBL; AF052293; AAC33564.1; JOINED.
 DR EMBL; AF052294; AAC33564.1; JOINED.
 DR EMBL; AF052295; AAC33564.1; JOINED.
 DR EMBL; AF052296; AAC33564.1; JOINED.
 DR EMBL; AF052297; AAC33564.1; JOINED.
 DR EMBL; AF052298; AAC33564.1; JOINED.
 DR EMBL; AF052299; AAC33564.1; JOINED.
 DR EMBL; AF052300; AAC33564.1; JOINED.
 DR EMBL; AF052301; AAC33564.1; JOINED.
 DR EMBL; AF052302; AAC33564.1; JOINED.
 DR EMBL; AF052303; AAC33564.1; JOINED.
 DR EMBL; AF052304; AAC33564.1; JOINED.
 DR EMBL; AF052305; AAC33564.1; JOINED.
 DR EMBL; AF052306; AAC33564.1; JOINED.
 DR EMBL; AF052307; AAC33564.1; JOINED.
 DR EMBL; AF052308; AAC33564.1; JOINED.
 DR EMBL; AF052309; AAC33564.1; JOINED.
 DR EMBL; AF052310; AAC33564.1; JOINED.
 DR EMBL; AF052311; AAC33564.1; JOINED.
 DR EMBL; AF052312; AAC33564.1; JOINED.
 DR EMBL; AF052313; AAC33564.1; JOINED.
 DR EMBL; AF052314; AAC33564.1; JOINED.
 DR EMBL; AF052315; AAC33564.1; JOINED.
 DR EMBL; AF052316; AAC33564.1; JOINED.
 DR EMBL; AF052317; AAC33564.1; JOINED.
 DR EMBL; AF052318; AAC33564.1; JOINED.
 DR EMBL; AF052319; AAC33564.1; JOINED.
 DR EMBL; AF052320; AAC33564.1; JOINED.
 DR EMBL; AF052321; AAC33564.1; JOINED.
 DR EMBL; AF052322; AAC33564.1; JOINED.
 DR EMBL; AF052323; AAC33564.1; JOINED.
 DR EMBL; AF052324; AAC33564.1; JOINED.
 DR EMBL; AF052325; AAC33564.1; JOINED.
 DR EMBL; AF052326; AAC33564.1; JOINED.
 DR EMBL; AF052327; AAC33564.1; JOINED.
 DR EMBL; AF052328; AAC33564.1; JOINED.
 DR EMBL; AF052329; AAC33564.1; JOINED.
 DR EMBL; AF052330; AAC33564.1; JOINED.
 DR EMBL; AF052331; AAC33564.1; JOINED.
 DR EMBL; AF052332; AAC33564.1; JOINED.
 DR EMBL; AF052333; AAC33564.1; JOINED.
 DR EMBL; AF052334; AAC33564.1; JOINED.
 DR EMBL; AF052335; AAC33564.1; JOINED.
 DR EMBL; AF052336; AAC33564.1; JOINED.
 DR EMBL; AF052337; AAC33564.1; JOINED.
 DR EMBL; AF052338; AAC33564.1; JOINED.
 DR EMBL; AF052339; AAC33564.1; JOINED.
 DR EMBL; AF052340; AAC33564.1; JOINED.
 DR EMBL; AF052341; AAC33564.1; JOINED.
 DR EMBL; AF052342; AAC33564.1; JOINED.
 DR EMBL; AF052343; AAC33564.1; JOINED.
 DR EMBL; AF052344; AAC33564.1; JOINED.
 DR EMBL; AF052345; AAC33564.1; JOINED.
 DR EMBL; AF052346; AAC33564.1; JOINED.
 DR EMBL; AF052347; AAC33564.1; JOINED.
 DR EMBL; AF052348; AAC33564.1; JOINED.
 DR EMBL; AF052349; AAC33564.1; JOINED.
 DR EMBL; AF052350; AAC33564.1; JOINED.
 DR EMBL; AF052351; AAC33564.1; JOINED.
 DR EMBL; AF052352; AAC33564.1; JOINED.
 DR EMBL; AF052353; AAC33564.1; JOINED.
 DR EMBL; AF052354; AAC33564.1; JOINED.
 DR EMBL; AF052355; AAC33564.1; JOINED.
 DR EMBL; AF052356; AAC33564.1; JOINED.
 DR EMBL; AF052357; AAC33564.1; JOINED.
 DR EMBL; AF052358; AAC33564.1; JOINED.
 DR EMBL; AF052359; AAC33564.1; JOINED.
 DR EMBL; AF052360; AAC33564.1; JOINED.
 DR EMBL; AF052361; AAC33564.1; JOINED.
 DR EMBL; AF052362; AAC33564.1; JOINED.
 DR EMBL; AF052363; AAC33564.1; JOINED.
 DR EMBL; AF052364; AAC33564.1; JOINED.
 DR EMBL; AF052365; AAC33564.1; JOINED.
 DR EMBL; AF052366; AAC33564.1; JOINED.
 DR EMBL; AF052367; AAC33564.1; JOINED.
 DR EMBL; AF052368; AAC33564.1; JOINED.
 DR EMBL; AF052369; AAC33564.1; JOINED.
 DR EMBL; AF052370; AAC33564.1; JOINED.
 DR EMBL; AF052371; AAC33564.1; JOINED.
 DR EMBL; AF052372; AAC33564.1; JOINED.
 DR EMBL; AF052373; AAC33564.1; JOINED.
 DR EMBL; AF052374; AAC33564.1; JOINED.
 DR EMBL; AF052375; AAC33564.1; JOINED.
 DR EMBL; AF052376; AAC33564.1; JOINED.
 DR EMBL; AF052377; AAC33564.1; JOINED.
 DR EMBL; AF052378; AAC33564.1; JOINED.
 DR EMBL; AF052379; AAC33564.1; JOINED.
 DR EMBL; AF052380; AAC33564.1; JOINED.
 DR EMBL; AF052381; AAC33564.1; JOINED.
 DR EMBL; AF052382; AAC33564.1; JOINED.
 DR EMBL; AF052383; AAC33564.1; JOINED.
 DR EMBL; AF052384; AAC33564.1; JOINED.
 DR EMBL; AF052385; AAC33564.1; JOINED.
 DR EMBL; AF052386; AAC33564.1; JOINED.
 DR EMBL; AF052387; AAC33564.1; JOINED.
 DR EMBL; AF052388; AAC33564.1; JOINED.
 DR EMBL; AF052389; AAC33564.1; JOINED.
 DR EMBL; AF052390; AAC33564.1; JOINED.
 DR EMBL; AF052391; AAC33564.1; JOINED.
 DR EMBL; AF052392; AAC33564.1; JOINED.
 DR EMBL; AF052393; AAC33564.1; JOINED.
 DR EMBL; AF052394; AAC33564.1; JOINED.
 DR EMBL; AF052395; AAC33564.1; JOINED.
 DR EMBL; AF052396; AAC33564.1; JOINED.
 DR EMBL; AF052397; AAC33564.1; JOINED.
 DR EMBL; AF052398; AAC33564.1; JOINED.
 DR EMBL; AF052399; AAC33564.1; JOINED.
 DR EMBL; AF052400; AAC33564.1; JOINED.
 DR EMBL; AF052401; AAC33564.1; JOINED.
 DR EMBL; AF052402; AAC33564.1; JOINED.
 DR EMBL; AF052403; AAC33564.1; JOINED.
 DR EMBL; AF052404; AAC33564.1; JOINED.
 DR EMBL; AF052405; AAC33564.1; JOINED.
 DR EMBL; AF052406; AAC33564.1; JOINED.
 DR EMBL; AF052407; AAC33564.1; JOINED.
 DR EMBL; AF052408; AAC33564.1; JOINED.
 DR EMBL; AF052409; AAC33564.1; JOINED.
 DR EMBL; AF052410; AAC33564.1; JOINED.
 DR EMBL; AF052411; AAC33564.1; JOINED.
 DR EMBL; AF052412; AAC33564.1; JOINED.
 DR EMBL; AF052413; AAC33564.1; JOINED.
 DR EMBL; AF052414; AAC33564.1; JOINED.
 DR EMBL; AF052415; AAC33564.1; JOINED.
 DR EMBL; AF052416; AAC33564.1; JOINED.
 DR EMBL; AF052417; AAC33564.1; JOINED.
 DR EMBL; AF052418; AAC33564.1; JOINED.
 DR EMBL; AF052419; AAC33564.1; JOINED.
 DR EMBL; AF052420; AAC33564.1; JOINED.
 DR EMBL; AF052421; AAC33564.1; JOINED.
 DR EMBL; AF052422; AAC33564.1; JOINED.
 DR EMBL; AF052423; AAC33564.1; JOINED.
 DR EMBL; AF052424; AAC33564.1; JOINED.
 DR EMBL; AF052425; AAC33564.1; JOINED.
 DR EMBL; AF052426; AAC33564.1; JOINED.
 DR EMBL; AF052427; AAC33564.1; JOINED.
 DR EMBL; AF052428; AAC33564.1; JOINED.
 DR EMBL; AF052429; AAC33564.1; JOINED.
 DR EMBL; AF052430; AAC33564.1; JOINED.
 DR EMBL; AF052431; AAC33564.1; JOINED.
 DR EMBL; AF052432; AAC33564.1; JOINED.
 DR EMBL; AF052433; AAC33564.1; JOINED.
 DR EMBL; AF052434; AAC33564.1; JOINED.
 DR EMBL; AF052435; AAC33564.1; JOINED.
 DR EMBL; AF052436; AAC33564.1; JOINED.
 DR EMBL; AF052437; AAC33564.1; JOINED.
 DR EMBL; AF052438; AAC33564.1; JOINED.
 DR EMBL; AF052439; AAC33564.1; JOINED.
 DR EMBL; AF052440; AAC33564.1; JOINED.
 DR EMBL; AF052441; AAC33564.1; JOINED.
 DR EMBL; AF052442; AAC33564.1; JOINED.
 DR EMBL; AF052443; AAC33564.1; JOINED.
 DR EMBL; AF052444; AAC33564.1; JOINED.
 DR EMBL; AF052445; AAC33564.1; JOINED.
 DR EMBL; AF052446; AAC33564.1; JOINED.
 DR EMBL; AF052447; AAC33564.1; JOINED.
 DR EMBL; AF052448; AAC33564.1; JOINED.
 DR EMBL; AF052449; AAC33564.1; JOINED.
 DR EMBL; AF052450; AAC33564.1; JOINED.
 DR EMBL; AF052451; AAC33564.1; JOINED.
 DR EMBL; AF052452; AAC33564.1; JOINED.
 DR EMBL; AF052453; AAC33564.1; JOINED.
 DR EMBL; AF052454; AAC33564.1; JOINED.
 DR EMBL; AF052455; AAC33564.1; JOINED.
 DR EMBL; AF052456; AAC33564.1; JOINED.
 DR EMBL; AF052457; AAC33564.1; JOINED.
 DR EMBL; AF052458; AAC33564.1; JOINED.
 DR EMBL; AF052459; AAC33564.1; JOINED.
 DR EMBL; AF052460; AAC33564.1; JOINED.
 DR EMBL; AF052461; AAC33564.1; JOINED.
 DR EMBL; AF052462; AAC33564.1; JOINED.
 DR EMBL; AF052463; AAC33564.1; JOINED.
 DR EMBL; AF052464; AAC33564.1; JOINED.
 DR EMBL; AF052465; AAC33564.1; JOINED.
 DR EMBL; AF052466; AAC33564.1; JOINED.
 DR EMBL; AF052467; AAC33564.1; JOINED.
 DR EMBL; AF052468; AAC33564.1; JOINED.
 DR EMBL; AF052469; AAC33564.1; JOINED.
 DR EMBL; AF052470; AAC33564.1; JOINED.
 DR EMBL; AF052471; AAC33564.1; JOINED.
 DR EMBL; AF052472; AAC33564.1; JOINED.
 DR EMBL; AF052473; AAC33564.1; JOINED.
 DR EMBL; AF052474; AAC33564.1; JOINED.
 DR EMBL; AF052475; AAC33564.1; JOINED.
 DR EMBL; AF052476; AAC33564.1; JOINED.
 DR EMBL; AF052477; AAC33564.1; JOINED.
 DR EMBL; AF052478; AAC33564.1; JOINED.
 DR EMBL; AF052479; AAC33564.1; JOINED.
 DR EMBL; AF052480; AAC33564.1; JOINED.
 DR EMBL; AF052481; AAC33564.1; JOINED.
 DR EMBL; AF052482; AAC33564.1; JOINED.
 DR EMBL; AF052483; AAC33564.1; JOINED.
 DR EMBL; AF052484; AAC33564.1; JOINED.
 DR EMBL; AF052485; AAC33564.1; JOINED.
 DR EMBL; AF052486; AAC33564.1; JOINED.
 DR EMBL; AF052487; AAC33564.1; JOINED.
 DR EMBL; AF052488; AAC33564.1; JOINED.
 DR EMBL; AF052489; AAC33564.1; JOINED.
 DR EMBL; AF052490; AAC33564.1; JOINED.
 DR EMBL; AF052491; AAC33564.1; JOINED.
 DR EMBL; AF052492; AAC33564.1; JOINED.
 DR EMBL; AF052493; AAC33564.1; JOINED.
 DR EMBL; AF052494; AAC33564.1; JOINED.
 DR EMBL; AF052495; AAC33564.1; JOINED.
 DR EMBL; AF052496; AAC33564.1; JOINED.
 DR EMBL; AF052497; AAC33564.1; JOINED.
 DR EMBL; AF052498; AAC33564.1; JOINED.
 DR EMBL; AF052499; AAC33564.1; JOINED.
 DR EMBL; AF052500; AAC33564.1; JOINED.
 DR EMBL; AF052501; AAC33564.1; JOINED.
 DR EMBL; AF052502; AAC33564.1; JOINED.
 DR EMBL; AF052503; AAC33564.1; JOINED.
 DR EMBL; AF052504; AAC33564.1; JOINED.
 DR EMBL; AF052505; AAC33564.1; JOINED.
 DR EMBL; AF052506; AAC33564.1; JOINED.
 DR EMBL; AF052507; AAC33564.1; JOINED.
 DR EMBL; AF052508; AAC33564.1; JOINED.
 DR EMBL; AF052509; AAC33564.1; JOINED.
 DR EMBL; AF052510; AAC33564.1; JOINED.
 DR EMBL; AF052511; AAC33564.1; JOINED.
 DR EMBL; AF052512; AAC33564.1; JOINED.
 DR EMBL; AF052513; AAC33564.1; JOINED.
 DR EMBL; AF052514; AAC33564.1; JOINED.
 DR EMBL; AF052515; AAC33564.1; JOINED.
 DR EMBL; AF052516; AAC33564.1; JOINED.
 DR EMBL; AF052517; AAC33564.1; JOINED.
 DR EMBL; AF052518; AAC33564.1; JOINED.
 DR EMBL; AF052519; AAC33564.1; JOINED.
 DR EMBL; AF052520; AAC33564.1; JOINED.
 DR EMBL; AF052521; AAC33564.1; JOINED.
 DR EMBL; AF052522; AAC33564.1; JOINED.
 DR EMBL; AF052523; AAC33564.1; JOINED.
 DR EMBL; AF052524; AAC33564.1; JOINED.
 DR EMBL; AF052525; AAC33564.1; JOINED.
 DR EMBL; AF052526; AAC33564.1; JOINED.
 DR EMBL; AF052527; AAC33564.1; JOINED.
 DR EMBL; AF052528; AAC33564.1; JOINED.
 DR EMBL; AF052529; AAC33564.1; JOINED.
 DR EMBL; AF052530; AAC33564.1; JOINED.
 DR EMBL; AF052531; AAC33564.1; JOINED.
 DR EMBL; AF052532; AAC33564.1; JOINED.
 DR EMBL; AF052533; AAC33564.1; JOINED.
 DR EMBL; AF052534; AAC33564.1; JOINED.
 DR EMBL; AF052535; AAC33564.1; JOINED.
 DR EMBL; AF052536; AAC33564.1; JOINED.
 DR EMBL; AF052537; AAC33564.1; JOINED.
 DR EMBL; AF052538; AAC33564.1; JOINED.
 DR EMBL; AF052539; AAC33564.1; JOINED.
 DR EMBL; AF052540; AAC33564.1; JOINED.
 DR EMBL; AF052541; AAC33564.1; JOINED.
 DR EMBL; AF052542; AAC33564.1; JOINED.
 DR EMBL; AF052543; AAC33564.1; JOINED.
 DR EMBL; AF052544; AAC33564.1; JOINED.
 DR EMBL; AF052545; AAC33564.1; JOINED.
 DR EMBL; AF052546; AAC33564.1; JOINED.
 DR EMBL; AF052547; AAC33564.1; JOINED.
 DR EMBL; AF052548; AAC33564.1; JOINED.
 DR EMBL; AF052549; AAC33564.1; JOINED.
 DR EMBL; AF052550; AAC33564.1; JOINED.
 DR EMBL; AF052551; AAC33564.1; JOINED.
 DR EMBL; AF052552; AAC33564.1; JOINED.
 DR EMBL; AF052553; AAC33564.1; JOINED.
 DR EMBL; AF052554; AAC33564.1; JOINED.
 DR EMBL; AF052555; AAC33564.1; JOINED.
 DR EMBL; AF052556; AAC33564.1; JOINED.
 DR EMBL; AF052557; AAC33564.1; JOINED.
 DR EMBL; AF052558; AAC33564.1; JOINED.
 DR EMBL; AF052559; AAC33564.1; JOINED.
 DR EMBL; AF052560; AAC33564.1; JOINED.
 DR EMBL; AF052561; AAC33564.1; JOINED.
 DR EMBL; AF052562; AAC33564.1; JOINED.
 DR EMBL; AF052563; AAC33564.1; JOINED.
 DR EMBL; AF052564; AAC33564.1; JOINED.
 DR EMBL; AF052565; AAC33564.1; JOINED.
 DR EMBL; AF052566; AAC33564.1; JOINED.
 DR EMBL; AF052567; AAC33564.1; JOINED.
 DR EMBL; AF052568; AAC33564.1; JOINED.
 DR EMBL; AF052569; AAC33564.1; JOINED.
 DR EMBL; AF052570; AAC33564.1; JOINED.
 DR EMBL; AF052571; AAC33564.1; JOINED.
 DR EMBL; AF052572; AAC33564.1; JOINED.
 DR EMBL; AF052573; AAC33564.1; JOINED.
 DR EMBL; AF052574; AAC33564.1; JOINED.
 DR EMBL; AF052575; AAC33564.1; JOINED.
 DR EMBL; AF052576; AAC33564.1; JOINED.
 DR EMBL; AF052577; AAC33564.1; JOINED.
 DR EMBL; AF052578; AAC33564.1; JOINED.
 DR EMBL; AF052579; AAC33564.1; JOINED.
 DR EMBL; AF052580; AAC33564.1; JOINED.
 DR EMBL; AF052581; AAC33564.1; JOINED.
 DR EMBL; AF052582; AAC33564.1; JOINED.
 DR EMBL; AF052583; AAC33564.1; JOINED.
 DR EMBL; AF052584; AAC33564.1; JOINED.
 DR EMBL; AF052585; AAC33564.1; JOINED.
 DR EMBL; AF052586; AAC33564.1; JOINED.
 DR EMBL; AF052587; AAC33564.1; JOINED.
 DR EMBL; AF052588; AAC33564.1; JOINED.
 DR EMBL; AF052589; AAC33564.1; JOINED.
 DR EMBL; AF052590; AAC33564.1; JOINED.
 DR EMBL; AF052591; AAC33564.1; JOINED.
 DR EMBL; AF052592; AAC33564.1; JOINED.
 DR EMBL; AF052593; AAC33564.1; JOINED.
 DR EMBL; AF052594; AAC33564.1; JOINED.
 DR EMBL; AF052595; AAC33564.1; JOINED.
 DR EMBL; AF052596; AAC33564.1; JOINED.
 DR EMBL; AF052597; AAC33564.1; JOINED.
 DR EMBL; AF052598; AAC33564.1; JOINED.
 DR EMBL; AF052599; AAC33564.1; JOINED.
 DR EMBL; AF052600; AAC33564.1; JOINED.
 DR EMBL; AF052601; AAC33564.1; JOINED.
 DR EMBL; AF052602; AAC33564.1; JOINED.
 DR EMBL; AF052603; AAC33564.1; JOINED.
 DR EMBL; AF052604; AAC33564.1; JOINED.
 DR EMBL; AF052605; AAC33564.1; JOINED.
 DR EMBL; AF052606; AAC33564.1; JOINED.
 DR EMBL; AF052607; AAC33564.1; JOINED.
 DR EMBL; AF052608; AAC33564.1; JOINED.
 DR EMBL; AF052609; AAC33564.1; JOINED.
 DR EMBL; AF052610; AAC33564.1; JOINED.
 DR EMBL; AF052611; AAC33564.1; JOINED.
 DR EMBL; AF052612; AAC33564.1; JOINED.
 DR EMBL; AF052613; AAC33564.1; JOINED.
 DR EMBL; AF052614; AAC33564.1; JOINED.
 DR EMBL; AF052615; AAC33564.1; JOINED.
 DR EMBL; AF052616; AAC33564.1; JOINED.
 DR EMBL; AF052617; AAC33564.1; JOINED.
 DR EMBL; AF052618; AAC33564.1; JOINED.
 DR EMBL; AF052619; AAC33564.1; JOINED.
 DR EMBL; AF052620; AAC33564.1; JOINED.
 DR EMBL; AF052621; AAC33564.1; JOINED.
 DR EMBL; AF052622; AAC33564.1; JOINED.
 DR EMBL; AF052623; AAC33564.1; JOINED.
 DR EMBL; AF052624; AAC33564.1; JOINED.
 DR EMBL; AF052625; AAC33564

KW Coiled coil; Muscle protein; Thick filament; Myosin; Phosphorylation.
 FT DOMAIN 1 47 NONHELICAL REGION (POTENTIAL).
 FT DOMAIN 48 871 COILED COIL (POTENTIAL).
 FT DOMAIN 872 882 NONHELICAL REGION (POTENTIAL).
 FT DISULFID 143 143 INTERCHAIN (POTENTIAL).
 FT DISULFID 632 632 INTERCHAIN (POTENTIAL).
 SO SEQUENCE 882 AA; 101949 MW; 894C4BFF4A670C5F CRC64;

Query Match 14.0%: Score 150.5; DB 1; Length 882;
 Best Local Similarity 24.4%: Pred. No. 0.12; Mismatches 83; Indels 63; Gaps 10;
 Matches 63; Conservative 49; Mismatches 83; Indels 63; Gaps 10;

QY 6 KLFFDLAEEENVDAEFLKNEKLSVKAQLS-----QKDKREKRS----- 45
 DB 447 KLHBLHAKKALADANKLHLEENKRLAGEIRELOTALKENDADQKRDENKQALRA 506
 QY 46 --QAIIDTLRDLERNATVESLONALN-----KADMLCSTLKK----- 82
 DB 507 ELQALRIEMERRLOEKEEEMELKRNLOFEIDRLIALADAEARKKSEI--SRUKKRYQA 564
 QY 83 QMKFLEQRQDFTKQAREEAHRLKCKMKTMEOIEL--LQSORSEVEEMIMDMVGOSAV 139
 DB 565 EIAELEMTVDNLNRANIEAQ--KTKRKSEQLKTLQASLEPTQQLQGVLDQYALAKRV 622
 QY 140 EQLAVYCVSLKKEYENLKEARK-----ATGELAD--RLKDKLVSSRSKLTLMTELD 189
 DB 623 AALSLELECTALDNMARRKQAEVDELEANGRISDLSINNLTSTKNLE---TFLS 679
 QY 190 QAKLELSAQKDLQSDQ 207
 DB 680 TAQADLDEVTELHADE 697

RESULT 30
 SCPI_HUMAN STANDARD: PRT: 976 AA.
 ID SCPI_HUMAN 015431; 014963;
 AC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Synaptonemal complex protein 1 (SCP-1 protein).
 GN SCPI OR SCPI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RX MEDLINE=97224467; PubMed=9119375;
 RA Meuwissen R.L.J., Meerts I., Hoovers J.M.N., Leschot N.J.,
 RA Heyting C.,
 RT "Human synaptonemal complex protein 1 (SCP1): isolation and
 RT characterization of the cDNA and chromosomal localization of the
 RT gene";
 RL Genomics 39:377-384(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RX MEDLINE=98037449; PubMed=9371398;
 RA Kondoh N., Nishina Y., Tsuchida J., Koga M., Tanaka H., Uchida K.,
 RA Iizawa J., Taketo M., Nozaki M., Nojima H., Matsumiya K., Namiki M.,
 RA Okuyama A., Nishimune Y.;
 RT "Assignment of synaptonemal complex protein 1 (SCP1) to human
 RT chromosome 1p13 by fluorescence in situ hybridization and its
 RT expression in the testis";
 RL Cytogenet. Cell Genet. 78:103-104(1997).
 CC -1- FUNCTION: MAJOR COMPONENT OF THE TRANSVERSE FILAMENTS OF
 CC SYNAPTONEMAL COMPLEXES (SCS). FORMED BETWEEN HOMOLOGOUS
 CC CHROMOSOMES DURING MEIOTIC PROPHASE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. IN TRIPARTITE SEGMENTS OF
 CC SYNAPTONEMAL COMPLEXES, BETWEEN LATERAL ELEMENTS IN THE NUCLEUS.

CC FOUND ONLY WHERE THE CHROMOSOME CORES ARE SYNAPSED. ITS N-TERMINUS
 CC IS FOUND TOWARDS THE CENTRE OF THE SYNAPTONEMAL COMPLEX WHILE THE
 CC C-TERMINUS EXTENDS WELL INTO THE LATERAL DOMAIN OF THE
 CC SYNAPTONEMAL COMPLEX (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: TESTIS.
 CC -1- DOMAIN: CONSISTS OF AN ALPHA-HELICAL STRETCH OF 700 AA RESIDUES,
 CC FLANKED BY N- AND C-TERMINAL GLOBULAR DOMAINS. THE C-TERMINAL
 CC DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X95654; CA64956.1; -;
 DR EMBL: D67035; BAA22586.1; -;
 DR MIM: 602162; -;
 KW Nuclear protein; Meiosis; Cell division; Phosphorylation;
 DB DNA-binding; Coiled coil.

FT DOMAIN 12 100 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 107 798 COILED COIL (POTENTIAL).
 FT DOMAIN 117 120 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 679 682 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 880 883 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 961 969 ARG/LYS-RICH (BASIC).
 FT CONFLICT 46 106 L->F (IN REF. 2).
 FT CONFLICT 106 106 F->Y (IN REF. 2).
 FT CONFLICT 153 153 K->C (IN REF. 2).
 FT CONFLICT 161 161 F->T (IN REF. 2).
 FT CONFLICT 168 168 E->D (IN REF. 2).
 FT CONFLICT 216 216 N->S (IN REF. 2).
 FT CONFLICT 225 226 HG->FE (IN REF. 2).
 FT CONFLICT 350 350 K->N (IN REF. 2).
 FT CONFLICT 360 360 E->D (IN REF. 2).
 FT CONFLICT 400 401 KN->NY (IN REF. 2).
 FT CONFLICT 406 406 K->I (IN REF. 2).
 FT CONFLICT 415 415 K->T (IN REF. 2).
 FT CONFLICT 449 449 E->D (IN REF. 2).
 FT CONFLICT 483 510 IQLVATTSQYVSKEVKDKLTLENEK->YSGHYHKW
 FT CONFLICT 516 528 TVLPKRGQRPKLSKRE (IN REF. 2).
 FT CONFLICT 516 528 LTVSHCNKLSLENK->YFTLQOASPPPN (IN REF. 2).
 FT CONFLICT 549 549 N->I (IN REF. 2).
 FT CONFLICT 560 560 K->T (IN REF. 2).
 FT CONFLICT 805 805 E->D (IN REF. 2).
 FT CONFLICT 941 941 P->S (IN REF. 2).
 SO SEQUENCE 976 AA; 114069 MW; 8BA81D042AC2696B CRC64;

Query Match 14.0%: Score 150. DB 1; Length 976;
 Best Local Similarity 26.2%: Pred. No. 0.14; Mismatches 86; Indels 28; Gaps 9;
 Matches 59; Conservative 52; Mismatches 86; Indels 28; Gaps 9;

QY 10 DLAAEEENVDAEFLKNEKLSVKAQLS-----QKDKREKRS----- 45
 DB 501 DLKTELEN-----EKLKTELTSHCNKLSLENKELTQETSDMTLEKNOQEDINNKKOEE 556
 QY 68 NALNKAEMLCSTLKKQKFLQRODFTKQAREEAHRLKCKMKTMEO---TELLQORS 123
 DB 557 RMLQIINLOETETQLNNELEYVAEELKQKRD---VKCKLDSEKNCNMLRQVEKNK 613
 QY 124 EVEEM-----IRPMYGQSAVEQLAVYCVSLKKEYENLKEARKATGELADRLKDLVSS 177
 DB 614 YIEELQENKALKKKGTAE--KOLNYETIKVNLLELELESAAKQKFEIDTYQKTEIDK 671
 QY 178 RSKTKTLNTELDQKL---ELRSQKDLQSDQ-----EISLRKK 215
 DB 672 KISEENLLEVEKAKVIADEVAVKLQKTDKRCQKHAEMVALMEK 716


```

RESULT 31
MYH8_HUMAN STANDARD: PRT; 1937 AA.
ID MYH8_HUMAN
AC P13535; O14910;
DT 01-JAN-1990 (Rel. 13, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, perinatal (MyHC-perinatal).
OS MYH8.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90323631; PubMed=2373371;
RA Karsch-Mizrachi I., Feghali R., Shows T.B. Jr., Leinwand L.A.;
RT "Generation of a full-length human perinatal myosin heavy-chain-
RL encoding cDNA.";
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95324556; PubMed=7601129;
RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,
RA Stegman H.H., Rubinstein N.A.;
RT "Characterization of a human perinatal myosin heavy-chain
RL transcript.";
RN [3]
RP SEQUENCE OF 502-1937 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90235862; PubMed=1691980;
RA Rober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RA Arnold H.H.;
RT "Identification of three developmentally controlled isoforms of human
RL myosin heavy chains.";
RN [4]
RP SEQUENCE OF 860-1937 FROM N.A.
RX MEDLINE=89234168; PubMed=2715179;
RA Feghali R., Leinwand L.A.;
RT "Molecular genetic characterization of a developmentally regulated
RL human perinatal myosin heavy chain.";
RN [5]
RP SEQUENCE OF 1-46 FROM N.A.
RA Esser K., Tidhar A., Myszkowski M.;
RT "Isolation and characterization of the human perinatal MHC promoter.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entries requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M36769; AAC17185.1; -
DR EMBL; Z38183; CAA86293.1; -
DR EMBL; X51592; CAA35941.1; -
DR EMBL; M35250; AAA36346.1; -
DR EMBL; AF067143; AAC621557.1; -
DR PIR; A30220; A30220.
DR HSSP; P08799; ILVK.
DR MIM; 160741; -.
DR InterPro; IPR000048; IQ.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Multigene family;
KW Calmodulin-binding.
FT DOMAIN 1 780
FT MYOSIN HEAD-LIKE.
FT DOMAIN 781 813
FT IQ.
FT DOMAIN 842 1937
FT COILED COIL (POTENTIAL).
FT NP_BIND 181 188
FT ATP.
FT DOMAIN 658 680
FT ACTIN-BINDING.
FT DOMAIN 760 774
FT ACTIN-BINDING.
FT MOD_RES 132 132
FT METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 698 698
FT ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 708 708
FT ALKYLATION (SH-2) (POTENTIAL).
FT CONFLICT 15 15
FT A -> R (IN REF. 2).
FT CONFLICT 970 970
FT E -> N (IN REF. 1 AND 4).
FT CONFLICT 1072 1072
FT M -> N (IN REF. 3).
FT CONFLICT 1247 1247
FT N -> H (IN REF. 1 AND 4).
FT CONFLICT 1251 1251
FT MC -> DG (IN REF. 3).
FT CONFLICT 1261 1261
FT E -> G (IN REF. 1 AND 4).
FT CONFLICT 1297 1297
FT K -> Q (IN REF. 1 AND 4).
FT CONFLICT 1377 1378
FT KY -> NT (IN REF. 3).
FT CONFLICT 1504 1505
FT EN -> AH (IN REF. 1 AND 4).
FT CONFLICT 1847 1847
FT E -> D (IN REF. 1 AND 4).
FT CONFLICT 1914 1914
FT D -> H (IN REF. 2).
SQ SEQUENCE 1937 AA; 222762 MW; A3EE2D151792E9E8 CRC64;
Query Match 14.0%; Score 150; DB 1; Length 1937;
Best Local Similarity 24.3%; Pred. No. 0.29;
Matches 64; Conservative 48; Mismatches 85; Indels 66; Gaps 9;
3 INKLFEEDLAEEENVLDAEFLKNELDVSAVQALQSDREKRSQALIDTLRDTLEENRAT 62
1272 LINDL---TRQRRRLQTEAGEYISROLDKDALVSQLSRSQASTQOIEELKHOLEETRA 1328
63 VESLQNLAKEMKSTLKKOMKFL-----OROE 93
1329 KNLALHQLSSRRDCLLRQYEEDEGKAELQRLASKANSEVAQWRKYETDAIQRTTE 1388
94 TKAREE-AHRLK-----CKRKTQETILLQSQSEVEEMIRPMGVQSAVEOL 142
1389 LEEAKKKLAQRLQEAEEHVEAVNAKCALEKTKQRLQ---NEVEDLMD-----VERS 1438
143 AVYCVLKEVEYENLKARKATGELADRLKKDVLSSRSKLTLTLELDQAK-----L 193
1439 NACCAALDKQRNFDFVLSWMKKYETQAELEASQKESNSTELFKYKNVYESLDL 1498
194 E-LRSAQKDLQSDQETSLRRK 215
1499 ETLRENNKIQ---GEISDLTEQ 1518

```


CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D89992; BAA22069.1; -
 DR EMBL: D50476; BAA09069.1; -
 DR EMBL: DA3700; BAA07802.1; -
 DR HSSP: P08799; 1AMD, IQ.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS00936; IQ; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Calmodulin-binding;
 KW Multigene family.
 KM DOMAIN 1 781 MYOSIN HEAD-LIKE.
 FT DOMAIN 1 782 IQ.
 FT DOMAIN 812 839 HINGE.
 FT DOMAIN 840 1935 COILED COIL (POTENTIAL).
 FT NP_BIND 178 185 ATP (POTENTIAL).
 FT DOMAIN 659 681 ACTIN-BINDING.
 FT DOMAIN 761 775 ACTIN-BINDING.
 FT MOD_RES 129 129 METHYLATION (SH-1).
 FT MOD_RES 699 699 ALKYLATION (SH-2).
 FT MOD_RES 709 709 ALKYLATION (SH-2).
 SQ SEQUENCE 1935 AA; 221599 MW; 9A1244B67D63C83B CRC64;

Query Match 13.98; Score 149; DB 1; Length 1935;
 Best Local Similarity 24.38; Pred. No. 0.33;

Matches 66; Conservative 52; Mismatches 84; Indels 70; Gaps 13;

OY 11 LAQEEENVLAELFLKNEELS-VKAKLSQKDR-----KRDSQAITDLRLDLE-----DRNA 61
 DB 1552 LEHRESITILVQLELNOYKSEIDRKLAKEDEBQIKRNSQRYVDSMSTLDSVRSRND 1611
 OY 62 TV---ESLQNALNKAEMLCSTLKKQ---MKFLERODETRQAR---EFAHR---LKCK 107
 DB 1612 ALRVKKKMEGDLNEMETIQLSHANQAAEAQKQLRNVOGQLKDAQLHDEAVRQGEDMKEQ 1671
 OY 108 MKIMEQJLELLQSORSEVEEMIRDMVGOSAVEQLA-----YCSLKRKY 153
 DB 1672 VAVVERNSLMOQAEIEELRALDEQTERGRVAVDELDASERVGLILHSQNTSLINTKKL 1721
 OY 154 E-----NLKEARKA-----TGELADRLKKDYSS-----RSK-----L 181
 DB 1732 EADLVQVQGEVDVAVQARNAAEKAKAITDAAMAAEELKKEQDTSAHLERMKKNLEVTY 1791
 OY 182 KTLNTELDQAK-LELSAQKQLQSDAQDETSL 212
 DB 1792 KDLQHRLEDAEASLAKMGKKQLQKLESVRREL 1823

RESULT 34
 MYH9_HUMAN STANDARD; PRT; 1960 AA.
 ID MYH9_HUMAN
 AC P35579; O60805;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Myosin heavy chain, nonmuscle type A (Cellular myosin heavy chain,
 type A) (Nonmuscle myosin heavy chain-A) (NMHC-A).

GN MYH9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
 RA Clamp M., Smink L.J., Alnsough R., Almeida J.P., Babbage A.,
 RA Baguley C., Bailey J., Barlow K., Bates K.N., Beasley O., Bird C.P.,
 RA Blakey S., Bridgeman A.M., Buck D., Burgess J., Burrill W.D.,
 RA Burton J., Carder C., Carter N.P., Chen Y., Clark G., Clegg S.M.,
 RA Cobley V., Cole C.G., Collier R.E., Connor R.E., Conroy D., Corby N.,
 RA Coville G.J., Cox A.V., Davis J., Dawson E., Dhani P.D., Dockree C.,
 RA Dodswoth S.J., Durbin R.M., Ellington A., Evans K.L., Fey J.M.,
 RA Fleming K., French L., Garner A.A., Gilbert J.G.R., Goward M.E.,
 RA Graffham D., Griffiths M.N., Hall C., Hall R., Hall-Tamlyn G.,
 RA Heathcote R.W., Ho S., Holmes S., Hunt S.E., Jones M.C., Kershaw J.,
 RA Kimberley A., King A., Laird G.K., Langford C.F., Leversha M.A.,
 RA Lloyd C., Lloyd D.M., Martyn I.D., Mashreghi Mohammad M.,
 RA Matthews L., Mccann O.T., Mcclellan J., McLaren S., Mcmurtry A.A.,
 RA Milne S.A., Mortimore B.J., Odell C.N., Pavitt R., Pearce A.V.,
 RA Pearson D., Phillimore B.J., Phillips S.H., Plumb R.W., Ramsay H.,
 RA Ramsey Y., Rogers L., Ross M.T., Scott C.E., Senra H.K., Skuce C.D.,
 RA Smalley S., Smith M.L., Soderlund C., Spragon L., Stewart C.A.,
 RA Sulston J.E., Swann R.M., Vandin M., Wall M., Wallis J.M.,
 RA Whiteley M.N., Willey D., Williams L., Williams S., Williamson H.,
 RA Wilmer T.E., Wilming L., Wright C.L., Hubbard T., Bentley D.R.,
 RA Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T.,
 RA Asakawa S., Kudoh J., Shintani A., Shibuya K., Yoshizaki Y., Aoki N.,
 RA Mitsuyama S., Roe B.A., Chen F., Chu L., Crabtree J., Descamps S.,
 RA Do A., Do T., Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S.,
 RA Lai H., Lao H.T., Lewis J., Lewis S., Lin S.P., Lon P., Malat E.,
 RA Nguyen T., Pan H., Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S.,
 RA Sloan D., Song L., Wang Q., Wang Y., Wang Z., White J., Willingham D.,
 RA Wu H., Yao Z., Zhan M., Zhang G., Chissoe S., Murray J., Miller N.,
 RA Minx P., Fulton R., Johnson D., Bents G., Bentley D., Bradshaw H.,
 RA Bourne S., Cordes M., Du Z., Fulton L., Goele D., Graves T.,
 RA Hawkins J., Hinds K., Kemp K., Latreille P., Layman D., Ozersky P.,
 RA Hocking T., Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K.,
 RA Nelson J., Korf I., Bedell J.A., Hillier L., Marais E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Salta S.,
 RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Beilman L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
 RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tiliahun Y., Wright H.;
 RL "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 RN [2]
 RP SEQUENCE OF 1-1337 FROM N.A.
 RX MEDLINE=92003925; PubMed=1912569;
 RA Toothaker L.E., Gonzalez D.A., Tung N., Lemons R.S., le Beau M.M.,
 RA Arnaout M.A., Clayton L.K., Tenen D.G.;
 RT "Cellular myosin heavy chain in human leukocytes: isolation of 5'
 RT cDNA clones, characterization of the protein, chromosomal
 RT localization, and upregulation during myeloid differentiation.";
 RL Blood 78:1826-1833(1991).
 RN [3]
 RP SEQUENCE OF 1-715 FROM N.A.
 RX MEDLINE=91316803; PubMed=1860190;
 RA Simons M., Wang M., McBride O.W., Kawamoto S., Yamakawa K.,
 RA Gula D., Adelsstein R.S., Weir L.;
 RT "Human nonmuscle myosin heavy chains are encoded by two genes located
 RT on different chromosomes.";
 RL Circ. Res. 69:530-539(1991).
 RN [4]
 RP SEQUENCE OF 714-1960 FROM N.A.
 RX MEDLINE=90138958; PubMed=1967836;
 RA Saez C.G., Myers J.C., Shows T.B., Leinwand L.A.;
 RT "Human nonmuscle myosin heavy chain mRNA: generation of diversity
 RT through alternative polyadenylation.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:1164-1168(1990).
RN [5]
RP VARIANT DFNA17 HIS-705.
RX MEDLINE-20489856; PubMed-11023810;
RA Lalwani A.K., Goldstein J.A., Kelley M.J., Luxford W., Castelein C.M.,
RA Mhatre A.N.,
RT "Human nonsyndromic hereditary deafness DFNA17 is due to a mutation in
RT nonmuscle myosin MYH9.";
RL Am. J. Hum. Genet. 67:1121-1128(2000).
RN [6]
RP VARIANTS MHA/FTNS/SBS K-93; C-702; C-1165; H-1424 AND K-1841.
RX MEDLINE-20428192; PubMed-10973259;
RA Seri M., Cusano M., Gangarossa S., Caridi G., Bordo D., Lo Nigro C.,
RA Iolagueri G.M., Ravazzolo R., Savino M., Del Vecchio M., d'Apolito M.,
RA Iolascon A., Zeltante L.L., Savoia A., Baldini C.L., Norris P.,
RA Magrini U., Belletti S., Heath K.E., Babcock M., Gluckman M.J.,
RA Alliprandis E., Bizzaro N., Desnick R.J., Martignetti J.A.,
RT "Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and
RT Sebastian syndromes.";
RL Nat. Genet. 26:103-105(2000).
RN [7]
RP VARIANTS MHA ILE-1155 AND LYS-1841.
RX MEDLINE-20428193; PubMed-10973260;
RA Kelley M.J., Jawien W., Ortel T.L., Korczak J.F.,
RT "Mutation of MYH9, encoding non-muscle myosin heavy chain A, in
RT May-Hegglin anomaly.";
RL Nat. Genet. 26:106-108(2000).
CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
CC CAPING.
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF MAY-HEGGLIN ANOMALY
CC (MHA), AN AUTOSOMAL, DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUCOCYTE INCLUSIONS.
CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF FECHTNER SYNDROME
CC (FTNS), AN AUTOSOMAL, DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
CC WITH THROMBOCYTOPENIA, GIANT PLATELETS AND LEUCOCYTE INCLUSIONS.
CC WITH ADDITIONAL ALPORT-LIKE CLINICAL FEATURES OF SENSORINEURAL
CC DEAFNESS, CATARACTS AND NEPHRITIS.
CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF SEBASTIAN SYNDROME
CC (SBS), AN AUTOSOMAL, DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUCOCYTE INCLUSIONS.
CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF AN AUTOSOMAL DOMINANT
CC FORM OF NONSYNDROMIC SENSORINEURAL DEAFNESS (DFNA17) WHICH IS
CC CHARACTERIZED BY PROGRESSIVE HEARING IMPAIRMENT AND
CC COCHLEOSACCULAR DEGENERATION.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: 282215; CAB05105.1; -
CC EMBL: M61105; AAA59888.1; -
CC EMBL: M69180; AAA61765.1; -
CC EMBL: M31013; AAA36349.1; -
CC HSSP: P08799; ILVK.
CC MIM: 160775; -
CC MIM: 153640; -
CC MIM: 155100; -
CC MIM: 603622; -
CC MIM: 605249; -
CC InterPro: IPR000048; IQ.

DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MISC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin: ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Alkylation; Multigene family; Disease mutation;
KW Deafness.
FT DOMAIN 1 778 MYOSIN HEAD-LIKE.
FT 779 808 IQ.
FT DOMAIN 837 1926 COILED COIL (POTENTIAL).
FT NP_BIND 174 181 ATP (POTENTIAL).
FT DOMAIN 654 676 ACTIN-BINDING.
FT MOD_RES 694 694 ALKYLTATION (SH-1) (POTENTIAL).
FT MOD_RES 704 704 ALKYLTATION (SH-2) (POTENTIAL).
FT VARIANT 93 93 N -> K (IN MHA).
FT VARIANT 93 93 /FTID-VAR.010791.
FT VARIANT 702 702 R -> C (IN FTNS).
FT VARIANT 705 705 R -> H (IN DFNA17).
FT VARIANT 1155 1155 /FTID-VAR.010793.
FT VARIANT 1165 1165 T -> I (IN MHA).
FT VARIANT 1165 1165 /FTID-VAR.010794.
FT VARIANT 1424 1424 R -> C (IN SBS).
FT VARIANT 1424 1424 /FTID-VAR.010795.
FT VARIANT 1841 1841 D -> H (IN FTNS).
FT VARIANT 1841 1841 /FTID-VAR.010796.
FT VARIANT 1841 1841 E -> K (IN MHA).
FT CONFLICT 53 55 /FTID-VAR.010797.
FT CONFLICT 660 660 EAL -> RGH (IN REF. 3).
FT CONFLICT 869 869 T -> S (IN REF. 3).
FT CONFLICT 931 931 C -> M (IN REF. 4).
FT CONFLICT 931 931 C -> Y (IN REF. 4).
FT CONFLICT 1240 1241 KG -> GR (IN REF. 4).
FT CONFLICT 1350 1350 E -> EE (IN REF. 2).
FT CONFLICT 1764 1764 T -> A (IN REF. 2).
FT CONFLICT 1771 1771 S -> G (IN REF. 2).
SQ SEQUENCE 1960 AA; 226531 MW; 588F84BB8C10656F CRC64;

Query Match 13.9%; Score 149; DB 1; Length 1960;
Best Local Similarity 19.7%; Pred. No. 0.33;
Matches 60; Conservative 68; Mismatches 87; Indels 90; Gaps 10;

QY 4 INKLPDLAEEENV-----LDAEF--LKNEIDSVKAOLSKR 38
DB 1071 IAEKQGLAKKEELDAALAVEEAAQKNMALKKTRELSQISELOEDLESERASRNKA 1130
QY 39 DREKRDQAIIIDRLPDLERNAFTESIALAKKAEMLCSTLKK-----QMKFLQG 89
DB 1131 EKKRDLGEELAEKLELEPDLSTAAQDLRSKRQDEVVITLKTTEEAKKTHEAOIQEM 1190
QY 90 RQ-----DETQAREAHRLKCKM---KTME-----QIELLOS-----ORSEV 125
DB 1191 RQHQSQAVERLADLBOLBQTKRVKANLEKAKQOTLENGELANEKVLLQGGDSEHRRKV 1250
QY 126 EEMIRDMGV-----GQSAVEDLAVYCVSLKKEYNLKARKATGLADRLKLDVSSRKV 181
DB 1251 EAQLOELQVKNFNGERVELADKVTYKLOVELDVTGLLSQSDSKSKSLTKDPSALESQI 1310
QY 182 K-----TLNTELDQAKLELSAOKDLOSADO-----EITSIR 213
DB 1311 QDTQELLOEENRQKLSLSTKLKQVDEKNSFRQOLEEBEBAKHLEKQIATTLHAQVADMK 1370
QY 214 KKSD 218

```

DB      1371 KKMED 1375

RESULT 35
MYH-CHICK STANDARD; PRT; 1978 AA.
ID MYH-CHICK STANDARD; PRT; 1978 AA.
AC P10587;
DT 01-JUL-1989 (Rel. 11, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, gizzard smooth muscle.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88118918; PubMed=2892941;
RA Yanagisawa M., Hamada Y., Katsuragawa Y., Imamura M., Mikawa T.,
RA Masaki T.,
RT "Complete primary structure of vertebrate smooth muscle myosin heavy
RT chain deduced from its complementary DNA sequence. Implications on
RT topography and function of myosin.";
RL J. Mol. Biol. 198;143-157(1987).
RN [2]
RP REVISIONS.
RA Masaki T.;
RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-203.
RX MEDLINE=88032919; PubMed=3312184;
RA Maier T., Onishi H., Yajima E., Matsuda G.;
RT "Amino acid sequence of the amino-terminal 24 kDa fragment of the
RT heavy chain of chicken gizzard myosin.";
RL J. Biochem. 102;133-145(1987).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF 1-818, AND REVISIONS 204-215.
RX MEDLINE=98412652; PubMed=9741621;
RA Dominguez R., Freyzo Y., Trybus K.M., Cohen C.;
RT "Crystal structure of a vertebrate smooth muscle myosin motor domain
RT and its complex with the essential light chain: visualization of the
RT pre-power stroke state.";
RL Cell 94;559-571(1998).
CC - FUNCTION: MUSCLE CONTRACTION.
CC - SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC - SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC - DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC - PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC - MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC - SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC - SIMILARITY: CONTAINS 1 IO DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X06546; CAA29793.1; -
DR PIR: S03166; S03166.
DR PDB: 1BR1; 09-SEP-98.

```

```

DR PDB: 1BR2; 09-SEP-98.
DR PDB: 1BR4; 09-SEP-98.
DR InterPro: IPR000048; IO.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IO; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IO; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IO; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Calmodulin-binding;
KW Multigene family; 3D-structure.
FT INIT MET 0 0
FT DOMAIN 1 790 MYOSIN HEAD-LIKE.
FT DOMAIN 791 820 IO.
FT DOMAIN 849 1978 RODLIKE TAIL (S2 AND LMW DOMAINS).
FT DOMAIN 849 1978 COILED COIL (POTENTIAL).
FT NP_BIND 176 183 ATP.
FT DOMAIN 666 688 ACTIN-BINDING.
FT DOMAIN 767 781 ACTIN-BINDING.
FT MOD_RES 1 127 BLOCKED.
FT MOD_RES 127 127 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 706 706 ALKYLATION (SH-1).
FT MOD_RES 716 716 ALKYLATION (SH-2).
FT CONFLICT 127 127 MISSING (IN REF. 3).
FT CONFLICT 204 215 KDTSTGGSFS -> RTFASLKVHLP (IN REF. 1).
SQ SEQUENCE 1978 AA; 228663 MW; B7B6C9235E273D93 CRC64;

```

```

Query Match 13.9%; Score 149; DB 1; Length 1978;
Best local similarity 24.3%; Pred. No. 0.33;
Matches 54; Conservative 48; Mismatches 86; Indels 34; Gaps 7;

OY 14 EENNVDAEF-----LNKELDSVKAQLSQDRKRSQAIIDLTLDLEERNATVESLQN 68
DB 892 EEKNLLQEKRLQAEITELTLYAEAEKRVRLAKKQE-----LEELHMEARIEEEE 941
OY 69 ALKKAEMLCSTLKKQKFFLEQRODERKQAREEHLKC-----KXKIMEQLELLQSO--- 121
DB 942 RSQQLQEKRRKQKQMDLEEEBAAQKQLQLEKVTADGKIKKEDDILIMEDQNNK 1001
OY 122 ---RSEVEEMIRDMGQSAVEQLAVYCVSLKKEYENLKEARATGELADRLKDLVSS 177
DB 1002 LTKERKLEERVDLTNNLAEEBEKAKNLTIKKNKHSM-----ISELEVRLLKE-EKS 1054
OY 178 RSKIKTLNTELDAQKLELRSQKDLQSDQDEI-TSLRKSD 218
DB 1055 ROELEKIKRLEGGSSDLHQIAELQAIQIAELAKLAQKKEE 1096

RESULT 36
MYSD_RABIT
ID MYSD_RABIT STANDARD; PRT; 501 AA.
AC Q99105;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Myosin heavy chain, embryonic smooth muscle isoform (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal aorta;
RX MEDLINE=91139672; PubMed=1995631;

```


